MODELING RELIABILITY GROWTH IN ACCELERATED STRESS TESTING

DISSERTATION

Jason K. Freels
Major, USAF

AFIT-ENS-DS-13-D-02

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT-ENS-DS-13-D-02

MODELING RELIABILITY GROWTH IN ACCELERATED STRESS TESTING

DISSERTATION

Presented to the Faculty

Department of Systems Engineering and Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy in Systems Engineering

Jason K. Freels

Major, USAF

December, 2013

AFIT-ENS-DS-13-D-02

MODELING RELIABILITY GROWTH IN ACCELERATED STRESS TESTING

Jason K. Freels
Major, USAF

Approved:

_____	_____
Dr. Joseph J. Pignatiello, USAF (Chair)	Date
_____	_____
Dr. Raymond R. Hill, USAF (Member)	Date
_____	_____
Dr. Richard L. Warr, Lt Col, USAF (Member)	Date
_____	_____
Dr. Joseph R. Wirthlin, Lt Col, USAF (Member)	Date
_____	_____
Dr. Adedeji B. Badiru Dean, Graduate School of Engineering and Management	Date

Abstract

Qualitative accelerated test methods improve system reliability by identifying and removing initial design flaws. However, schedule and cost constraints often preclude sufficient testing to generate a meaningful reliability estimate from the data obtained in these tests. In this dissertation a modified accelerated life test is proposed to assess the likelihood of attaining a reliability requirement based on tests of early system prototypes. Assuming each prototype contains an unknown number of independent competing failure modes whose respective times to occurrence are governed by a distinct Weibull law, the observed failure data from this qualitative test are shown to follow a poly-Weibull distribution.

However, using an agent-based Monte Carlo simulation, it is shown that for typical products subjected to qualitative testing, the failure observations result from a homogenous subset of the total number of latent failure modes and the failure data can be adequately modeled with a Weibull distribution. Thus, the projected system reliability after implementing corrective action to remove one or more failure modes can be estimated using established quantitative accelerated test data analysis methods. Our results suggest that a significant cost and time savings may be realized using the proposed method to signal the need to reassess a product's design or reallocate test resources to avoid unnecessary maintenance or redesigns. Further, the proposed approach allows a significant reduction in the test time and sample size required to estimate the risk of meeting a reliability requirement over current quantitative accelerated life test techniques.

Additional contributions include a numerical and analytical procedure for obtaining the maximum likelihood parameter estimates and observed Fisher information matrix components for the generalized poly-Weibull distribution. Using this procedure, we show that the poly-Weibull distribution outperforms the best-fit modified Weibull alternatives in the literature with respect to their fit of reference data sets for which the hazard rate functions are non-monotone.

To mom and dad

Acknowledgments

I would like to express my deep appreciation and gratitude to my advisor, Dr. Joseph Pignatiello, for the patient guidance and mentorship he provided to me, throughout the process of completing this degree. It is because of Dr. Pignatiello's investment that I have become a far better researcher and writer than when I began this effort. I am truly fortunate to have had the opportunity to work with him.

I would also like to thank my committee members, Drs. Raymond Hill, Richard Warr and Joseph Wirthlin for the friendly guidance, thought provoking suggestions, and the general collegiality that each of you offered to me over these three years. In a similar vein, I also want to recognize Dr. Som Soni for his contributions to my intellectual growth during my years of study at the Air Force Institute of Technology and for making this entire opportunity possible.

Finally, I'd be remiss if I didn't acknowledge the immeasurable sacrifices made by my wife in shouldering far more than her fair share of the family burdens while I pursued this final degree.

Jason K. Freels

Table of Contents

Abstract iv

Acknowledgments vi

List of Figures x

List of Tables xvi

List of Acronyms xvii

I. Introduction 1

Background 1

Problem Statement, Objectives and Scope 5

Dissertation Outline and Research Impacts 6

II. Literature Review 9

Introduction 9

Competing Risks Analysis 10

Overview 10

Statistical Notions and Notation 11

Parametric Competing Risks Analysis 15

Identifiability Paradox 16

Nonparametric Competing Risks 18

Reliability Growth with Competing Risks 21

Reliability Growth Modeling 23

Non-Accelerated Reliability Growth 23

Accelerated Reliability Growth. 26

Highly Accelerated Life Testing (HALT) 28

Step Stress Accelerated Life Testing Data Analysis 34
Conclusion 37
III. Bridging the Gap between Quantitative and Qualitative Accelerated Life Tests .. 38
Introduction 38
Highly Accelerated Life Testing 39
Example HALT Scenario and Data Structure 42
Modeling Approaches 45
Quantitative Accelerated Life Testing Model 45
Physics of Failure Analysis 47
Accelerated Reliability Growth 50
McLean's AFR Estimator Model 53
Future Research Avenues 55
Conclusion 56
IV. Maximum Likelihood Estimation for the Poly-Weibull Distribution 58
Introduction 58
The Poly-Weibull Distribution 60
Parameter Estimation 63
Application 64
Aarset data 65
Meeker data - Uncensored 68
Meeker data - Censored 71
Conclusion 73
V. Characterizing Reliability Growth in Early System Design 74
Introduction 74

Testing Framework 76
Model 77
Single Failure Mode Case – No Corrective Action 77
Single Failure Mode Case – With Corrective Action 83
Design of Accelerated Reliability Growth Tests 89
Multiple Failure Modes Case – With Corrective Action 94
Monte Carlo Simulation Procedure 97
Conclusions and Future Work 108
VI. Conclusions and Recommendations 110
Dissertation Summary 110
Conclusions and Future Work 113
Extending the Model to Multiple Corrective Action Periods 113
Extending the Model to Multiple Stressor Tests 116
Bibliography 117
Appendix A: Poly–Weibull Observed Fisher Information Equations 127
Appendix B: Histograms of Corrected Akaike Information Criterion Values $AICc$ for the Weibull, Bi-Weibull and Tri-Weibull Distributions 129
Appendix C: R Computer Code Used to Compare the Fit of Various Distributions to the Aarset and Meeker Datasets as Discussed in Chapter IV 190
Appendix D: R Computer Code Used to Simulate and Analyze Poly-Weibull Failure Observations as Discussed in Chapter V 207
Vita 220

List of Figures

Figure 1 – Illustration of idealized reliability growth planning curves plotted along with assessed growth across three test phases 4

Figure 2 – Depiction of competing failure modes and occurrence times using the David and Moeschberger notion 12

Figure 3 – Idealized reliability growth planning curves for three test phases showing the difference in planned and assessed reliability (Source: MIL-HDBK-189C, p. 11)24

Figure 4 – Illustration of product specification, operating and destruct limits before and after highly accelerated life testing 31

Figure 5 – Four-step highly accelerated life testing process 40

Figure 6 – Relationship between product operating and destruct limits 40

Figure 7 – Effect of highly accelerated life testing on bathtub curve 42

Figure 8 – Example step-stress profile 43

Figure 9 – Fraction of failures discovered by common HALT stressors (McLean, 2008) 44

Figure 10 – Idealized reliability growth planning curves developed across three test phases showing the difference in planned and assessed reliability (Source: MIL-HDBK-189C, p. 11) 51

Figure 11 – Example bi-Weibull density (top) and hazard (bottom) functions 61

Figure 12 – Example tri-Weibull density (top) and hazard (bottom) functions 62

Figure 13 – Triple plot illustrating the fit of the bi-Weibull, tri-Weibull NMW and EMWE models to the Aarset data 67

Figure 14 – Triple plot illustrating the fit of the bi-Weibull, tri-Weibull NMW and EMWE models to the uncensored Meeker data 70

Figure 15 – Triple plot illustrating the fit of the bi-Weibull, tri-Weibull NMW and EMWE models to the censored Meeker data 72

Figure 16 – Weibull plot of failure times observed at three levels of temperature in an example accelerated life test 80

Figure 17 – Weibull plot of example failure data redrawn parallel under the assumption that the shape parameter is independent of stress 80

Figure 18 – Arrhenius relationship plot of the example data..... 82

Figure 19 – Adjusted Arrhenius relationship plot of example data versus temperature 83

Figure 20 – Updated Arrhenius plot of example data after corrective action..... 86

Figure 21 – Joint plot of density and reliability functions at 180°C with corrective action 87

Figure 22 – Variance comparison of the two– and three–stress accelerated test designs with equal allocation against the optimal two-stress test design 92

Figure 23 – Comparison of the reduction in total test time for the two-stress equal apportionment design the two-stress optimal design and the three-stress equal apportionment design for various levels of extrapolation factor 93

Figure 24 – Serial arrangement of flaws within the prototypes subjected to qualitative accelerated reliability test. The arrangement demonstrates the competing risk assumption in the model where the time to failure for prototype i is the minimum activation time among the J flaws..... 95

Figure 25 – Example poly–Weibull hazard functions with various parameter values 96

Figure 26 – Graphical depiction of process to simulate observed minimum failure times for systems with varied numbers of competing failure modes and levels of initial design quality 99

Figure 27 – Relationship between the coefficient of variation and shape parameter β_j for Weibull distributed random variables 100

Figure 28 – Plots illustrating the distribution of the coefficient of variation values for multiple levels of initial design quality 101

Figure 29 – Results of the Anderson-Darling goodness of fit tests showing the fraction of tests with significance levels greater than 0.05 for each combination of quality level sample size and system complexity.....105

Figure 30 – Sigmoid shape of $d(S)$.....115

Figure 31 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 10$ systems, quality = low 130

Figure 32 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 20$ systems, quality = low 131

Figure 33 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 40$ systems, quality = low 132

Figure 34 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 100$ systems, quality = low 133

Figure 35 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 10$ systems, quality = mid 134

Figure 36 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 20$ systems, quality = mid 135

Figure 37 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions J = 5 modes, n = 40 systems, quality = mid 136

Figure 38 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 100$ systems, quality = mid 137

Figure 39 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 10$ systems, quality = high 138

Figure 40 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 20$ systems, quality = high 139

Figure 41 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 40$ systems, quality = high 140

Figure 42 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 5$ modes, $n = 100$ systems, quality = high 141

Figure 43 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 10$ systems, quality = low 142

Figure 44 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 20$ systems, quality = low 143

Figure 45 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 40$ systems, quality = low 144

Figure 46 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 100$ systems, quality $=$ low 145

Figure 47 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 10$ systems, quality $=$ mid 146

Figure 48 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 20$ systems, quality $=$ mid 147

Figure 49 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 40$ systems, quality $=$ mid 148

Figure 50 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 100$ systems, quality $=$ mid 149

Figure 51 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 10$ systems, quality $=$ high 150

Figure 52 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 20$ systems, quality $=$ high 151

Figure 53 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 40$ systems, quality $=$ high 152

Figure 54 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 10$ modes, $n = 100$ systems, quality $=$ high 153

Figure 55 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 20$ modes, $n = 10$ systems, quality $=$ low 154

Figure 56 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 20$ modes, $n = 20$ systems, quality $=$ low 155

Figure 57 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 20$ modes, $n = 40$ systems, quality $=$ low 156

Figure 58 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 20$ modes, $n = 100$ systems, quality $=$ low 157

Figure 59 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 20$ modes, $n = 10$ systems, quality $=$ mid 158

Figure 60 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions $J = 20$ modes, $n = 20$ systems, quality $=$ mid 159

Figure 61 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 20 modes, n = 40 systems, quality = mid 160

Figure 62 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 20 modes, n = 100 systems, quality = mid 161

Figure 63 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 20 modes, n = 10 systems, quality = high 162

Figure 64 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 20 modes, n = 20 systems, quality = high 163

Figure 65 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 20 modes, n = 40 systems, quality = high 164

Figure 66 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 20 modes, n = 100 systems, quality = high 165

Figure 67 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 10 systems, quality = low 166

Figure 68 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 20 systems, quality = low 167

Figure 69 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 40 systems, quality = low 168

Figure 70 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 100 systems, quality = low 169

Figure 71 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 10 systems, quality = mid 170

Figure 72 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 20 systems, quality = mid 171

Figure 73 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 40 systems, quality = mid 172

Figure 74 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 100 systems, quality = mid 173

Figure 75 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 10 systems, quality = high 174

Figure 76 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 20 systems, quality = high 175

Figure 77 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 40 systems, quality = high 176

Figure 78 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 40 modes, n = 100 systems, quality = high 177

Figure 79 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 10 systems, quality = low 178

Figure 80 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 20 systems, quality = low 179

Figure 81 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 40 systems, quality = low 180

Figure 82 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 100 systems, quality = low 181

Figure 83 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 10 systems, quality = mid 182

Figure 84 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 20 systems, quality = mid 183

Figure 85 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 40 systems, quality = mid 184

Figure 86 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 100 systems, quality = mid 185

Figure 87 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 10 systems, quality = high 186

Figure 88 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 20 systems, quality = high 187

Figure 89 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 40 systems, quality = high 188

Figure 90 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions J = 50 modes, n = 100 systems, quality = high 189

List of Tables

Table 1 – HALT stress sequence 29
Table 2 – McLean HALT stress regimen 41
Table 3 – Physics of failure process 49
Table 4 – AFR estimator Inputs 54
Table 5 – The Aarset data set 65
Table 6 – Maximum likelihood estimates (standard errors) for the Aarset data 66
Table 7 – Performance measures for the Aarset data set 66
Table 8 – The Meeker data set 68
Table 9 – Maximum likelihood estimates (standard errors) for the uncensored Meeker data set 69
Table 10 – Performance measures for the uncensored Meeker data set 69
Table 11 – Maximum likelihood estimates for the censored Meeker data set 71
Table 12 – Performance measures for the censored Meeker data set 71
Table 13 – Projected reliability measures at 180^{o}C after corrective action 85

List of Acronyms

Acronym	Definition
AF	acceleration factor
AFR	actual failure rate
AIC	Akaike information criterion
ALT	accelerated life testing
AMSAA	Army Materiel Sustainment and Analysis Activity
ARG	accelerated reliability growth
BIC	Bayesian information criterion
CALT	calibrated accelerated life testing
CDF	cumulative density function
CIF	cumulative incidence function
CoV	coefficient of variation
DL	destruct limit
DoF	degree of freedom
ERS	engineered resilient systems
EWME	exponentiated modified Weibull extension distribution
FEF	fix effectiveness factor
FLT	fundamental limit of technology
HALT	highly accelerated life testing
KM	Kaplan-Meier
KS	Kolmogorov-Smirnov test
MEOST	multi-environment over-stress testing
MLE	maximum likelihood function
MTBF	mean time between failures
NHPP	non-homogenous Poisson process
NMW	new modified Weibull distribution
OL	operating limit
PLP	power law process
PoF	physics of failure
PPoF	probabilistic physics of failure
QALT	quantitative accelerated life testing
QAST	qualitative accelerated stress testing
ROCOF	rate of occurrence of failures
SSALT	step-stress accelerated life testing
TAAF	test-analyze-and-fix
TFR	tempered failure rate
TFT	test-fix-test

MODELING RELIABILITY GROWTH IN ACCELERATED STRESS TESTING

I. Introduction

> *"Our record of predicting where we will use military force since Vietnam is perfect – we have never once gotten it right, there isn't a single instance: Grenada, Panama, the first Gulf War, the Balkans, Haiti, you can just keep going through the list, where we knew and planned for such a conflict six months in advance."*
>
> - *Secretary of Defense Robert Gates, May 24, 2011*

Background

Recapitalizing an aging Air Force inventory requires a balanced approach in which neither major combat operations against near-peer technology nor operations such as those employed in Afghanistan and Iraq are over-emphasized [1]. With this in mind, then Secretary of Defense, Robert Gates [2], unveiled seven strategic priorities to focus future science and technology investments and guide the development of next generation of systems engineering tools and processes. According to Neches [3], the Secretary's second priority, Engineered Resilient Systems (ERS), comprises efforts to "efficiently create, field and evolve systems which can readily adapt to the inevitable changes in threat, technology and mission environments." Under the ERS construct, future systems must be robust to a wide range of possible operational environments, not just a single known scenario. An implication of this strategy is the need for critical components capable of "plug and play" operations across platforms and operating environments.

Recent history, however, indicates that US defense systems do not satisfy their operational suitability requirements sufficiently often [4]. Consequently, the Department of Defense spends too much for system redesigns, spares management, and maintenance. Reports commissioned to investigate the root causes of these reliability and suitability shortfalls [5, 6] identified a lack of appropriate systems engineering processes, specifically a robust reliability growth strategy beginning early in the development cycle as a primary contributor. In response to these reports, the Reliability Improvement Working Group adopted [7, 8] to align Department of Defense policy with the best practices of reliability management, and to provide the most value with the least risk in terms of fielding reliable products. These standards specify a scientific approach to design and build reliability into products early-on and institutionalize the creation of a comprehensive reliability growth strategy throughout the acquisition cycle. As result of their implementation a greater burden is placed on verifying the maturity of early designs, thereby minimizing the expenditure of test resources in subsequent development phases.

Reliability growth is the positive improvement in a product's reliability distribution parameter over a period of time due to changes in product design or manufacturing processes [9]. Reliability growth modeling has historically played a role in determining whether major development efforts, such as military weapon systems, are likely to meet reliability requirements in time for graduation to the next development phase, and eventually to operational testing. To assess product reliability, prototypes are subjected to a series of tests which exercise the system

using a selected subset of inputs from the overall set of possible inputs that may be encountered during the product's lifetime. Inputs may represent the expected usage environment or comprise a specific stress profile directed by the customer.

Introducing a complex product often requires a lengthy development process during which it is expected that the reliability will steadily improve based on testing, failure mode discovery, root-cause analysis and design changes or component substitutions. Testing may be composed of many different types of tests, each with its own objectives. Developmental tests identify the technical capabilities and limitations of proposed designs and ensure sufficient design maturity is achieved prior to operational testing. In operational testing the focus is on demonstrating that the design is suitable for its intended use in a realistic operational environment. Entering operational test with an immature design often results in continued debugging into the early life of the product after it has been released to the market, usually at much greater cost than if the fault were discovered in developmental testing. Therefore it is desirable to model the improvement in reliability over time to (1) forecast the length of the development process, (2) ensure proper allocation of testing resources and (3) estimate the reliability upon entry into market.

Ideally, the assessment of reliability growth should begin soon after program initiation with the development of an idealized reliability growth planning curve (Figure 1). Once testing begins, progress can be gauged by comparing quantitative assessments of system reliability against the idealized curve. Close agreement between the planned and demonstrated reliability is indicative of a successful

reliability growth program. Assessments below the idealized curve indicate that reliability is lagging and may signal the need to reallocate test resources or reevaluate the testing strategy.

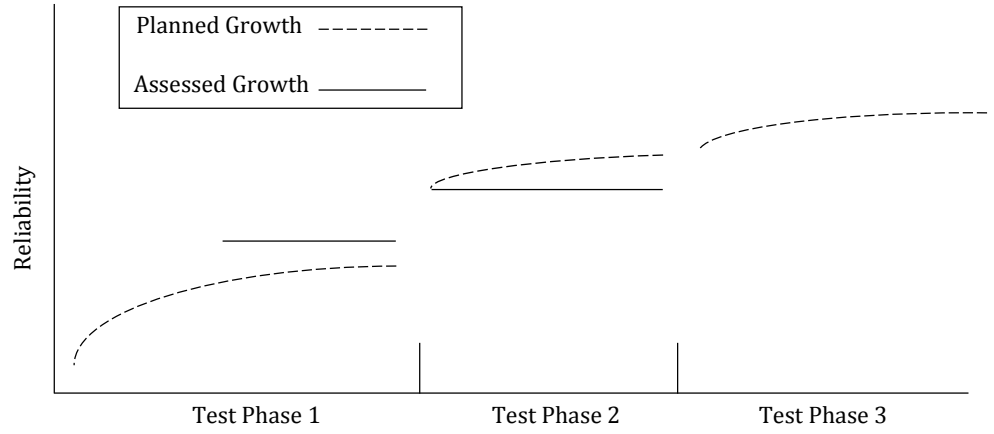

Figure 1 – Illustration of idealized reliability growth planning curves plotted along with assessed growth across three test phases

Many defense acquisition programs, however, fail to get on their reliability growth planning curves due to the existence of too many failure modes [4] at the start of developmental testing. Prior to formal testing it is assumed that few failure modes remain in the system for which the root cause has not been identified and understood. Mil-HDBK-189 [9] indicates that a key predictor of program success with respect to meeting the reliability goal through growth testing is that the initial system reliability be at least 30% of the required reliability. Achieving this level of initial reliability requires conducting design for reliability (DfR) activities to discover and remove both functional and operational failure modes [10] from early system prototypes. Functional failures occur when a latent design flaw activates to become a failure during normal operation. Operational failures result when the

operating environment to which the system is exposed falls outside of the expected normal range. With the availability of specialized materials and sophisticated manufacturing processes, functional failures have become rare for all but the most complex systems. Operational failures, however, are more difficult to predict and can result from poorly defined requirements, uncertain operating conditions and complex multi-stress interactions.

Qualitative accelerated stress testing (QAST) was developed to address both functional and operational failures by exposing products to elevated stresses and accelerate the identification and removal of these failure modes. QAST techniques, such as [11, 12, 13], employ a test-fix-test reliability improvement strategy to expand a product's operating limits and thereby ensure the highest reliability in the designed environment while potentially enabling usage in unplanned environments.

Problem Statement, Objectives and Scope

The most well-known QAST technique, highly accelerated life testing (HALT), utilizes a step-increasing stress profile to discover failure modes quickly. While published case-studies [14, 15, 16] demonstrate the effectiveness of HALT to improve product reliability, no published models exist to quantify the reliability improvement gained as a result of its use. Meeker [17] questions whether sufficient data are produced during HALT to construct a meaningful estimate of a system's reliability improvement. Typically, qualitative tests are conducted on very small samples of early system prototypes (typically 4 – 10) to iteratively find and remove failure modes at elevated stresses resulting in successively more mature designs.

The product improvement process used in HALT closely resembles a test-fix-test reliability growth test, therefore it is hypothesized that an early estimate of system reliability can be derived by modeling HALT as a reliability growth test.

In this dissertation a reliability growth projection model is developed to estimate system reliability as a result of implementing corrective actions to remove failure modes exposed in accelerated stress environments. The projection model is the first to utilize failure data obtained from qualitative accelerated tests and provides a statistically rigorous and defensible measure of the likelihood that a complex system or subsystem can attain its reliability requirement within an allocated test time. As a result of incorporating failure data obtained at elevated stresses into growth testing, the projection model enables a significant reduction in the number samples and the total test time required to estimate the reliability of the improved system. This ensures that the Department of Defense (DoD) resources allocated for testing can be used to address problems that may pose a significant risk to system performance after fielding. As consequence, the cost of implementing redesigns for fielded systems may also be avoided. This is especially important as DoD resources are likely to be reduced in the future.

Dissertation Outline and Research Impacts

This dissertation follows the scholarly article format where Chapters III, IV and V represent stand-alone papers that have either already been submitted for publication or will be submitted after graduation. These Chapters sequentially

develop the proposed model for assessing the reliability growth of a product with independent competing failure modes exposed to accelerated stresses.

The main contribution of this dissertation is the first model to assess the likelihood of attaining a reliability requirement through reliability growth testing where the test environment is not representative of the anticipated field operating conditions. A consequence of merging the accelerated testing and reliability growth methodologies is that one or more of the implicit assumptions associated with each individual methodology is violated. Chapter III, therefore, outlines several accelerated test data analysis techniques currently in use and illustrates why these techniques are insufficient to model data obtained from a qualitative accelerated stress test. A framework to guide future research efforts is then presented along with specific next steps detailing ways in which data from these qualitative tests may be incorporated into a reliability estimate for a complex system.

In Chapter IV, it is shown that for systems with independent competing failure modes whose respective times to occurrence are each governed by a distinct Weibull law, the observed system failure times follow a poly-Weibull distribution with vector valued parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. A numerical and analytical procedure is derived for obtaining the maximum likelihood parameter estimates and standard errors for the generalized poly-Weibull distribution with an arbitrary number of terms. The procedure is then used to show that the poly-Weibull distribution is capable of fitting data generated from complex failure processes with bathtub-

shaped hazard functions better than the best-fit modified Weibull alternatives in the literature.

Chapter V presents the testing methodology and associated reliability growth projection model used to estimate the likelihood of achieving a reliability requirement as result of qualitative accelerated stress testing. It is demonstrated, using published accelerated test data, that the proposed model can assess the reliability risk associated with critical systems or subsystems but with less than half of the samples and total test time required by current practices.

To frame the argument of accelerated reliability growth modeling, Chapter II presents an expanded literature review introducing significant concepts necessary for this research that may be unfamiliar to the reader. Topics discussed in this Chapter include reliability growth modeling, competing risks analysis, highly accelerated life testing and the analysis of step-stress accelerated life test data. Chapter VI concludes the dissertation, providing recommendations for further implementation of the model in design for reliability contexts and outlines future areas of research to improve the accuracy and robustness of the model.

II. Literature Review

"Six months in the lab will easily save you a half-day in the library."

- Ron Kerans

Introduction

To frame the argument of accelerated reliability growth modeling, this Chapter presents an expanded literature review introducing several major components pertaining to the research that may be unfamiliar to the reader. The first section provides an in-depth discussion on the analysis of time to event data when competing risks are present. Next, reliability growth modeling is discussed along with several concepts necessary to link traditional reliability growth modeling and accelerated testing to introduce accelerated reliability growth modeling. Then, an overview of highly accelerated life testing (HALT) is presented, with particular emphasis on how the test is conducted, the purpose of the test and how HALT differs from other well-known reliability testing techniques. Finally, current approaches to modeling time to failure in step-stress accelerated life testing are discussed specifically the cumulative damage model [18] and the tampered failure rate model [19].

Competing Risks Analysis

Overview

A competing risk is an event whose occurrence fundamentally alters or altogether eliminates the probability of observing an event of interest [20]. As an example, the probability of a woman developing breast cancer may become zero if she dies from another risk factor such as a heart attack or a stroke. Alternatively, the breast cancer risk for the same individual may be altered, positively or negatively, after receiving a diagnosis of lung cancer as the prescribed medical treatment could affect the mechanisms by which both types of cancer cells are created. Under the general competing risks assumption a system or individual is considered to be at risk of "failure" from J risks. The risks may be mutually independent or have some level of interdependence. David and Moeschberger [21] note, however, that where modeling the dependence among the risks was necessary, authors often grouped the risks into categories where independence among the categories could be assumed. Thus the assumption of mutual independence among the risks has dominated much of the competing risks literature.

Under a given set of conditions, each risk competes to be the cause of the failure, thus the term risk is reserved for an event yet to occur, while cause describes the particular risk from which the system actually failed. Furthermore, it is also assumed that systems subject to competing risks can fail from only one risk and can fail only once. It is tempting to consider the survival times for the remaining $J - 1$ risks that were not the cause of failure as randomly right-censored, however this

type of non-informative censoring scheme is not the correct approach in the presence of competing risks. In non-informative censoring, the only assumption made about the eventual occurrence of an event of interest is that there is some positive probability that the event will occur after the censoring event is observed. But in a competing risks framework once the system has failed due to risk $j = 1, \dots, J$ the probability of observing the system fail due to any of the remaining $J - 1$ risks is altered, and an informative censoring scheme is required. Putter et al. [22] analyzed competing risks data under both informative and non-informative censoring mechanisms, and showed that a non-informative analysis overestimates the true failure probability. Siannis et al. [23] developed a sensitivity model to measure the dependence between the lifetime of an individual and the censoring mechanism. Their results showed that the bias introduced by a small degree of dependence between the risks can have a noticeable effect on the analysis.

Statistical Notions and Notation

In the competing risks literature, two statistical notions, represented by David and Moeschberger [21] and Crowder [24] dominate the analytical methodology. Under the David and Moeschberger notion an increasing sequence of latent failure times is envisioned for each risk $j = 1, \dots, J$ assuming that j is the only risk to which the system is exposed (Figure. 2).

$$System \begin{cases} Mode_A \rightarrow A_1 \ \dots \ A_n \\ Mode_B \rightarrow B_1 \ \dots \ B_n \\ \dots \\ \dots \\ \dots \\ Mode_J \rightarrow J_1 \ \dots \ J_n. \end{cases}$$

Figure 2 – Depiction of competing failure modes and occurrence times using the David and Moeschberger notion

The notation of competing risks data according to David and Moeschberger is represented by the random variables, C_j and Z, where C_j is an indicator representing the risk which caused the failure and $Z = \min(T_1, T_2, \dots, T_j)$, where T_j is the system's theoretical lifetime assuming j is the only risk present. If $Z > t$, then

$$P\{Z > t\} = P\{T_1 > t, T_2 > t, \dots, T_j > t\} = R_Z(t) = 1 - F_Z(t) \qquad (1)$$

is the system reliability function. When mutual independence among the risks can be assumed, (1) becomes

$$R_Z(t) = \prod_{i=1}^{J} R_j(t) \qquad (2)$$

where $R_j(t)$ is the risk-specific reliability function for risk $j = 1, \dots, J$. Likewise, the system hazard rate function for mutually independent risks is

$$h_Z(t) = \frac{f_Z(t)}{R_Z(t)} = \sum_{i=1}^{J} h_j(t) \qquad (3)$$

where $h_j(t)$ is known as the cause-specific hazard rate function for risk $j = 1, \dots, J$ in the presence of competing risks. David and Moeschberger [21] further derive three

additional statistics commonly used in the field of demographics analysis that may have value in a reliability growth context. The net probaility of failure

$$q_j(t_1, t_2) = 1 - \exp\left[-\int_{t_1}^{t_2} r_j(t)dt\right] \tag{4}$$

measures the probability of failure for risk j assuming it is the only risk present. The crude probability of failure

$$Q_j(t_1, t_2) = \int_{t_1}^{t_2} g_j(x) \exp\left[-\int_{t_1}^{t} r_z(t)dt\right] dx \tag{5}$$

measures the probability of failure for risk j in the presence of other competing risks. Finally, the partial crude probability of failure

$$Q_{jk}(t_1, t_2) = \int_{t_1}^{t_2} g_j^{-k}(x) \exp\left[-\int_{t_1}^{t} r_z^{-k}(t)dt\right] dx \tag{6}$$

is the probability of failure for risk j assuming some of the competing risks have been eliminated where $g_j^{-k}(t)dt$ and $r_z^{-k}(t)$ are the conditional failure probability and hazard rate for cause j in the absence of cause k. Equations (4) – (6) are not limited to the assumption of independent risks, although when independence is assumed the equations can be simplified by replacing $g_j(t)$ with $r_j(t)$.

An alternative notion of competing risks it that of Crowder [24], who modeled competing risks data as a bivariate distribution of the time to failure, Z, and an indicator variable, C, representing the cause of failure. Crowder specifies the joint model in terms of the sub-distribution function $F(j,t) = P(C = j, T \leq t)$ or sub-survivor functions $R(j,t) = P(C = j, T > t)$ developed by Cox [25] where

$$F(j,t) + R(j,t) = q_j. \tag{7}$$

The sub-distribution function is not a proper distribution function as $\lim_{t\to\infty} F(j,t) = q_j$ rather than unity. Similarly, the sub-survivor function is not a proper reliability function in that $R(j,t) \neq P(T > t|C = j)$. The proper reliability function is instead given by $R(j,t)/q_j$ where the proportion $q_j > 0,\ j \in 1, \dots, J$ represents the marginal probability of the random variable C. Thus $q_j = P(C = j) = F(j,\infty) = R(j,0)$, subject to the constraint $\sum_{j=1}^{J} q_j = 1$. It follows that for failure times represented by the continuous random variable T the sub-density function, $f(j,t) = -dR(j,t)/dt$ and the respective marginal reliability and density functions are then $R(t) = \sum_{j=1}^{J} R(j,t)$ and $f(t) = \sum_{j=1}^{J} f(j,t)$. Finally, the sub-hazard function is expressed as $h(t,j) = f(t,j)/R(t)$ and the overall system hazard function is $h(t) = \sum_{j=1}^{J} h(t,j)$.

Equations similar to (4) –(6) have also been developed using the Crowder notation. Similar to, $q(t_1,t_2)$, the net probability of failure, Crowder represents the distribution of failure times caused by risk j as

$$U(t,j) = P(T = t|C = j) = f(j,t)/q_j. \tag{8}$$

Alternatively, the distribution of risks at a given age t is represented by

$$V(j,t) = (C = j|T = t) = f(j,t)/f(t). \tag{9}$$

Although $V(j,t)$ appears very different from the crude failure probability shown in (6), both functions serve similar functions. Ma and Krings [26] provided an excellent comparison of the David and Moeschberger [21] and Crowder [24] notions, and illustrated the conditions under which both are equivalent.

Parametric Competing Risks Analysis

When prior knowledge exists on either the risk-specific system lifetime distributions or the underlying failure time process, parametric approaches to competing risk analysis may be used. These techniques can reduce the analysis to exercises in multivariate statistics wherein the parameter values may be estimated using maximum likelihood or Bayesian estimation techniques. When no prior knowledge exists on the risk-specific distributions a nonparametric approach is required. These nonparametric approaches are discussed in the next section.

In reality, the risk-specific distributions are rarely known, thus much of parametric competing risks analysis is predicated on the concept of theoretical or latent failure times. As was described above, the latent failure times parametric approach presented by David and Moeschberger assumes that there exists an unobservable sequence of ordered failure times represented by the multivariate random variable $Y = \{Y_{11} \ldots Y_{1n}, \ldots, Y_{J1} \ldots Y_{Jn}\}$. For each failure mode j, the sequence $\{Y_{j1} \ldots Y_{jn}\}$ indicates the failure times that would be observed if j were the only mode in the system. In the presence of competing risks, the observable quantities are the minimum failure time $Z_i = \min[Y_{11}, Y_{21}, \ldots Y_{j1}]$ and the failure cause indicator C_i. Summarizing these quantities, the observed system lifetime conditioned on knowing the cause of failure is $X_{ij} = \left(Y_i \middle| Y_i = \min_j(Y_j)\right)$. Assuming independence among the risks, the pdf is then

$$f(x_i) = \frac{1}{P\left\{Y_i = \min_j Y_j\right\}} \cdot \left[f_{Y_i}(x)\right] \cdot \prod_{j=1, j\neq i}^{J} S_j(x). \tag{10}$$

Assuming that the number of items failed due to cause j is a random variable whose value can be represented by the multinomial distribution function $f(n_1, \ldots n_k) = \frac{n!}{\prod_{i=1}^{k} n_i!} \prod_{i=1}^{k} \left(P\left\{Y_i = \min_j Y_j\right\}\right)^{n_i}$, the likelihood function is then

$$L = \frac{n!}{\prod_i^k n_i!} \cdot \prod_{i=1}^{k} \left[\prod_{j=1}^{n} p_i(x_{ij}) \cdot \prod_{l=1, l\neq i}^{k} R_l(X_{ij})\right]. \tag{11}$$

Defining the bracketed term as L_i, equation (12) simplifies to

$$L = \frac{n!}{\prod_i^k n_i!} \prod_{i=1}^{k} L_i. \tag{12}$$

Equations (11) and (12) show that if the latent failure times for each risk follow a different distribution, parameter estimation is accomplished by maximizing each L_i term individually.

Identifiability Paradox

A disturbing complication exists in the latent failure times approach leading to a source of ongoing controversy in competing risks analysis known as the identifiability paradox. Cox [27] first noted the flaw in the parametric competing risks approach while discussing various models of failure times with two dependent risks. Tsiatis [28] later generalized the issue to p dependent risks, leading Crowder [24] to rename the issue as the Cox-Tsiatis Impasse – the issue is described as follows. As discussed in the beginning of this section, parametric competing risks

assumes a distribution of the latent failure times $\{Y_{1j}, Y_{2j} \ldots Y_{nj}\}$ for each failure mode j under the assumption that it is the only mode to which the system is subjected. However, since only the minimum failure times, $Z = \min[Y_{11}, Y_{12}, \ldots Y_{1j}]$ are actually observed, the true distributions of the latent failure times are unknown. Further, the distribution of observed lifetimes $Z_i, i = 1, 2, \ldots, I$ is completely determined by the joint distribution of the latent failure times Y_{ij}, however, the inverse is not necessarily true. Tsiatis [28] showed that for any joint reliability function with arbitrary dependence between the risks there exists another joint reliability function with independent risks that produces the exact same sub-density function $f(j,t)$. Therefore it is impossible to determine which model is correct as both will fit the data equally well. If the risks are independent no issue exists, otherwise, the statistical results may mislead the entire analysis.

Consider an example in which two dependent random variables Y_1 and Y_2 exist and represent the failure times for failure modes 1 and 2, respectively. Upon observing an i.i.d. sample $(Z, C_j)_{j=1,\ldots,J}$ Tsiatis [28] proved that for any distribution of (Z, C), there exist independent random variables Q_1 and Q_2 that provide the same distribution. This impasse has led subsequent research to focus on observable quantities rather than the joint distributions, thereby estimating the specific or "crude" hazard rate $\tilde{\lambda}_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(Z_j \in [t, t+\Delta t], C_j = 1 | Y > t)$, rather than the overall or latent hazard rate $\lambda_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(Z_j \in [t, t+\Delta t] | Z_j > t)$.

Nonparametric Competing Risks

Caplan et al. [29] broadly classifies failure events as either true or cause-specific failures with each classification having its own statistical methods that are not necessarily valid for the other. For true failures, the statistical methods are based on the fundamental assumption that all of the units under test would ultimately fail due to the failure mode of interest were the study allowed to continue for a sufficient amount of time. Methods developed to analyze such data assume the underlying cause for censoring observations is independent of the mechanism for the events occurrence. In other words, the survival time of an individual (or the time at which a subject experiences an event) is assumed to be independent of the mechanism that would cause the study to be censored. In theory, individuals for whom the observations are censored have an equal risk of event of interest compared to those still under study but have not yet observed the event. This is commonly referred to as non-informative censoring and is the basis of differentiation between the two commonly used non-parametric estimation techniques in competing risks analysis: the complement to the Kaplan-Meier estimator ($1 - KM$) and the Cumulative Incidence Function (CIF).

Nonparametric time-to-event curves are routinely presented in the literature with the Kaplan-Meier (KM) product limit estimator [30] being the most widely used. But this approach may not be appropriate when the analysis is focused on time-to-first event in scenarios where there are competing events. When there are no competing risks, the CIF and the $1 - KM$ estimators produce the same result. In

situations with competing risks, the cumulative incidence function is more appropriate. Cheng et al. [31] notes that in the literature the CIF has been referred to as the cause-specific risk, the crude incidence curve and the cause-specific failure probability. Additional references to CIF described below are consistent with the notation of Kalbfleisch and Prentice [32].

Gooley et al. [20] and Putter et al. [22] show that utilizing the $1 - KM$ estimator in the presence of competing risks tends to overestimate occurrence rate of each event. The discussion below compares the two methods, presupposing that a population of items is at risk of failure from two distinct failure modes, denoted below as mode 1 and mode 2. Further, the interest is in estimating the probability of failure due to failure mode 1, vice mode 2. Thus, competing risks are present, i.e. the hazard rate functions for both modes of failure exist and the number of failures from the competing risk will influence both the number of failures and the probability of failure due to the mode of interest. The hazard rate function

$$h(t) = \lim_{\Delta t \to 0} \frac{R(t) - R(t + \Delta t)}{\Delta t \times R(t)} \tag{13}$$

is a fundamental relation in the analysis of any time-to-event data. At time t the hazard rate is conditioned on the number of units at risk of failure, thus an estimator of $h(t)$ should be consistent with the simple ratio of the number of failures divided by the number of overall units under study. Recall that the Kaplan-Meier product limit estimator (KM) is a nonparametric maximum likelihood estimate of the reliability function $R(t)$. Conversely the complement to the Kaplan-

Meier estimator $1 - KM$ is a nonparametric estimate of the CDF. Using the notation of Kalbfleisch and Prentice [32] we define:

$n \equiv$ the initial number of items at risk
$f_i \equiv$ the number of items failed from the event of interest prior to time t_i
$r_i \equiv$ *the* number of items failed from the competing risk prior to time t_i
$n_i \equiv$ the number of items at risk after time t_i,

and the total number of units at risk of failure at any time t is expressed as $n_i = n - \sum_{t_i<t}(f_i + r_i)$. It follows that the respective Kaplan-Meier estimators for failure modes 1 and 2 are

$$KM_1(t) = \prod_{t_i<t}\left(1 - \frac{f_i}{n_i}\right) \quad \text{and} \quad KM_2(t) = \prod_{t_i<t}\left(1 - \frac{r_i}{n_i}\right). \tag{14}$$

which shows clearly that KM_1 is directly related to the hazard rate of the failure mode of interest but is independent of the competing failure mode. Thus, $KM_1(t)$ is not interpretable as the true probability of failure mode 1 when in the presence of failure mode 2.

Alternatively, the cumulative incidence function (CIF) estimator is a function of the hazard rates for both modes making CIF the preferred method of estimating the failure probability when competing risks are present. Kalbfleisch and Prentice [32] develop a cumulative incidence function from the overall Kaplan-Meier survivor function $KM_{12}(t) = KM_1(t) \times KM_2(t)$ expressed as

$$CIF(t) = \sum_{i=1}^{s} \frac{f_i}{n_i - 1} \times KM_{12}(t_i) \tag{15}$$

where s is the largest i such that $t_i < t$. Gooley et al. [20] notes that both $1 - KM_1$ and the cumulative incidence estimator in (15) are marginal estimates of the

probability of failure. Thus, both estimators will "jump" by an amount equal to n^{-1} after each occurrence of failure mode 1, until the first occurrence of failure mode 2. When failure mode 1 follows failure mode 2

$$\Delta_{1-KM_1} = \Delta_{CIF} \times (r_i/n_i). \quad (16)$$

For each subsequent occurrence of failure mode 2 the difference between the estimators continues to expand by r_i/n_i.

Reliability Growth with Competing Risks

Competing risks occur frequently in survivability and reliability analysis and a number of methods have been proposed for the analysis of this type of data. However, Ma and Krings [26] note that the interaction between competing risks analysis and reliability gradually withered during the period when significant advances were made in competing risks analysis. Consequently, the application of competing risks analysis in engineering reliability has fallen behind the theory of competing risks analysis.

Corcoran et al. [33] developed the first reliability growth projection model for estimating reliability after implementing corrective action in the final stage of development of an "expensive item." The Corcoran [33] reliability projection model is suitable for use in cases where corrective actions are installed at the conclusion of a single test phase consisting of *N* independent trials where the number of trial outcomes of interest is a multinomial random variable with parameters *N* (total number of trials), p_0 (unknown initial reliability), and q_i (unknown failure probability for failure mode $i = 1, \ldots, k$). Since a multinomial model is used, the

equality $p_0 + \sum_{i=1}^{k} q_i = 1$ must be satisfied, which imposes the restriction that at most one failure mode can occur per trial. In addition to deriving an exact expression for system reliability under the conditions above, seven different estimators are developed and evaluated for point estimation. These estimators were studied for bias, consistency, conservatism, etc. and ultimately it was shown that an unbiased estimate of the corrected system could not be obtained. The authors were the first to advance the idea of reducing initial failure probabilities by a fractional amount with consideration to fix effectiveness. Under their model, the expected reliability (under competing risks) at the end of the current test phase is given by

$$R(N|q_i) = R_I + \sum_{i}^{k} \rho_i q_i - \frac{1}{N} \sum_{i}^{k} \rho_i [1 - (1 - q_i)^N] \tag{17}$$

where N is the total number of failures, q_i is the failure probability of failure mode i, R_I is the initial reliability, and ρ_i is the fix effectiveness factor (FEF) for failure mode i. Dahiya [34] showed that six of the seven estimators initially considered by Corcoran et al. [33] possess the same limiting normal distribution allowing direct confidence interval and goodness of fit procedures for large samples. Olsen [35] showed how some of the estimators could be utilized under a multi-stage test program and developed a suitable variant of the Corcoran model.

Reliability Growth Modeling

Non-Accelerated Reliability Growth

Introducing a complex product often requires a lengthy development process during which it is expected that the reliability will steadily improve based on testing, failure mode discovery, root-cause analysis and design changes or component substitutions. Testing may be composed of many different types of tests, each with its own objectives. Developmental tests identify the technical capabilities and limitations of proposed designs and ensure sufficient design maturity is achieved prior to operational testing. In operational testing the focus is on demonstrating that the design is suitable for its intended use in a realistic operational environment. Entering operational test with an immature design often results in continued debugging into the early life of the product after it has been released to the market, usually at much greater cost than if the fault were discovered in developmental testing. Therefore, it is desirable to model the improvement in reliability over time to (1) forecast the length of the development process, (2) ensure proper allocation of testing resources and (3) estimate the reliability upon entry into market.

The initial product of a reliability growth model is the idealized reliability growth planning curve (Figure 3). Created early in the development process, the idealized curve is a roadmap to baseline the reliability growth progress within a single test phase. When testing continues across test phases, multiple idealized curves are developed. An indication that the reliability is lagging (assessments below the idealized curve) may signal the need to reallocate test resources or

reevaluate the testing strategy to ensure the reliability achieves the requirement with a specified level of confidence. Many models plot reliability growth as an increase in mean time between failures ($MTBF$), or a decrease in the failure rate ($MTBF^{-1}$) against cumulative test time across all units at risk [36].

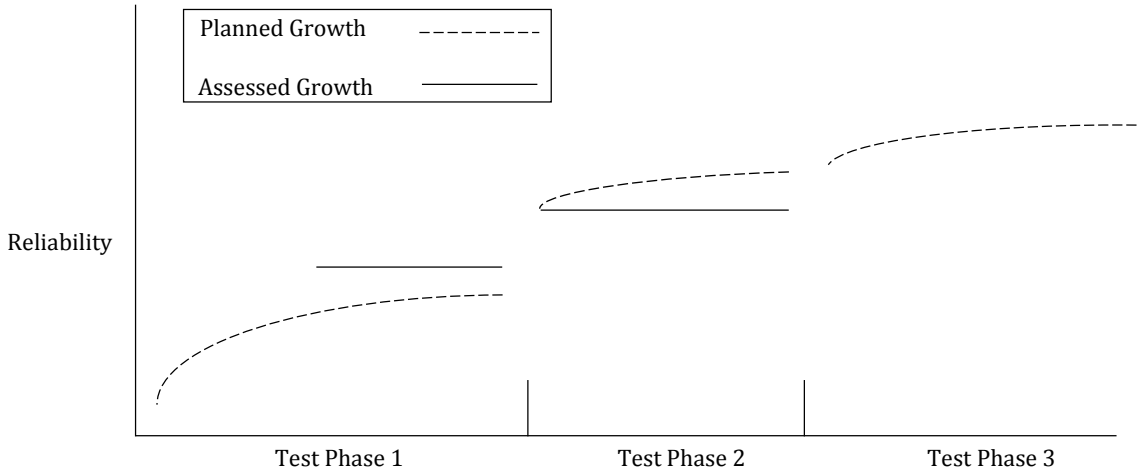

Figure 3 – Idealized reliability growth planning curves for three test phases showing the difference in planned and assessed reliability (Source: MIL-HDBK-189C, p. 11)

When many modular tests and debugging efforts are proceeding in parallel, growth may be measured as the rate at which defects are found and corrected in a testing interval. In principle, reliability growth models may also allow for decreases in reliability. For example, a product's reliability may be adversely affected as a result of substituting a cheaper but less reliable component. Our focus here, though, is on positive reliability growth which improves product reliability.

According to Ebeling [37] the earliest developed and most frequently used reliability growth model is that of Duane [38] who postulated from his empirical observations that for products under development, plots of cumulative failure rate

versus cumulative operating hours on a log-log scale were approximately linear with negative slope. Crow [39] derived the stochastic basis for the Duane model, referred to in the literature as a power law process (PLP) or informally as a Weibull process. The Crow [39] model is a type of nonhomogeneous Poisson process (NHPP) with a Weibull hazard intensity function $h(t) = \lambda\beta t^{\beta-1}$ where $\lambda > 0$ and $\beta > 0$ are interpreted, in the Crow model context, as the initial reliability of the system and the growth parameter, respectively. For $\beta < 1$, the hazard is decreasing (reliability is increasing) while $\beta > 1$ implies an increasing hazard and $\beta = 1$ means that both the reliability and the hazard rate are unchanged. Duane [38] observed β values in the neighborhood of 0.50 while the Army Materiel Sustainment Analysis Activity (AMSAA) [9] suggests values in the range (0.30, 0.75), depending on the level of commitment to reliability improvement and the type of system under development. The rigorous development of the PLP allows estimation and prediction in both the maximum likelihood estimation (MLE) and Bayesian contexts and has made the approach popular in the literature.

However, the PLP and other models based on the Duane postulate are an idealization of the true underlying Test-Analyze-And-Fix (TAAF) failure process. Sen [40] notes that these models estimate the improvement in reliability based on the number of failure modes discovered without accounting for fixes or design changes. Various reliability growth models have been developed since the Duane Postulate to plan, track, and project the reliability improvement of a system throughout the development process. Planning models are used to develop the

idealized and planned growth curves to address program schedules, required testing resources and the realism of achieving the reliability requirement with a desired level of confidence within the allocated test program. Tracking models gauge the progress of the reliability effort based on test results. Projection models indicate the anticipated reliability at some future time based on achievement to date and engineering assessments. Hall [10] provides one of the most detailed and comprehensive reviews of reliability growth planning, tracking and projection models for both continuous use and discrete use systems found anywhere in the literature. The review comprises a synopsis of over 80 papers covering planning models, tracking and projection models. Further, numerous reliability growth surveys/handbooks and thirty-six other papers covering theoretical results, simulation studies, real-world applications, personal perspectives, international standards, or related statistical procedures are also included.

Accelerated Reliability Growth

As opposed to the reliability growth models discussed by Hall [41] in which growth is based on a decreasing rate of failure mode discovery in successive test intervals, step-stress testing results in an increasing number of failure modes as testing progresses to higher stresses. Further, step-stress testing is a continuous process with no discernible intervals and the removal of each failure mode discovered does not yield the same growth at the non-accelerated use stress. Therefore, it is claimed that an improved estimate of the system's field reliability can be obtained by analyzing HALT failure data as a test-fix-test process and

distinguishing between the modes of failure that occur in a test through the incorporation of competing risks in the analysis.

Upon starting a traditional reliability growth test, it is assumed that few failure modes remain in the system for which the root cause has not been identified and understood. However, even when this assumption is met it is not uncommon for growth testing to require several months to attain the reliability goal with the required confidence level. Thus, accelerated test methods are increasingly utilized as part of a reliability growth strategy to reduce the overall test duration. Several widely used design improvement techniques [12, 42, 11, 13, 43] expose early product designs to accelerated environmental and repetitive use stresses to quickly force latent defects to manifest themselves thereby ensuring a highly reliable product is fielded quickly. The defining characteristic of these tests lies in the underlying assumption that simply meeting the validation requirement is insufficient for ensuring high field reliability as the requirement rarely encompasses the full stress envelope a product will actually encounter over its service life. Thus, in many instances failure modes exposed by these techniques are not representative of field experience, but reflect brief high-level stress states often not considered.

In a reliability growth test it is assumed that (1) a set of latent failure modes exists in each system under test, (2) the test profile will activate a sufficient number of these latent modes during the allocated test time and (3) corrective action removes a fraction of each discovered mode's failure intensity. In other words, if the mode specific failure rate due to mode i is represented by λ_i, the overall system

failure rate is $\sum_i \lambda_i$. If each failure mode receives corrective action that reduces λ_i by the proportion ρ, one would conclude that fixing any failure mode reduces the overall system failure rate by $\lambda_i(1-\rho)$.

Often the test profile involves the application of constant stresses chosen to represent field conditions. Thus the proper selection of stressors and stress levels is critical to achieving adequate reliability growth. If the stress profile is a function of time, however, consideration must be given to the stress level at which a failure mode is discovered. Under progressive stress or step-stress profiles it is possible to discover failure modes at high stress levels that would never be activated at the use stress level regardless of the exposure time. Thus, implementing corrective action on these failure modes may produce little to no actual reliability growth.

Highly Accelerated Life Testing (HALT)

Highly accelerated life testing originated in the electronics industry when Hobbs began to use the term in 1988 [44]. Though the mindset and process of HALT goes back several decades, Hobbs coined the term to describe a sequential testing process wherein products are subjected to a variety of stresses to identify the weak links in the design. More formally, HALT involves the application of individual or simultaneous stressors at levels elevated beyond those experienced in either the use environment or in traditional accelerated life tests. During testing, failure modes are discovered and removed from the design through corrective action in an iterative process which expands the product's operating margin. Once the test is complete, the product will have reduced random-failure probabilities and longer

lifetimes for wear-out in the designed usage environment. It should be noted, however that HALT does not specifically address wear-out failure modes.

According to McLean [12], HALT proceeds by first applying individual stressors that are increased step-wise at periodic intervals until a failure event occurs (i.e. no censoring). No standard exists to direct what stressors should be included for a given product type. However, McLean notes that experience has led practitioners to adopt an ordered set of stressors (Table 1) as standard to HALT which represents what is used in the majority of tests conducted to date. As the goal of HALT is to remove as many failure modes as possible, practitioners are not limited to the stressors listed in Table 1 nor are they required to ensure that each stressor is included. Product specific stresses are often included either independently or in combination with another test.

Table 1 – HALT stress sequence

Stressor	Notes	% Total Failure Modes
Cold Step Stress		14%
Hot Step Stress		17%
Rapid Thermal Transitions	$\geq 60\ C^0$/minute	4%
Vibration Step Stresses	6 DoF random vibrations	45%
Rapid Thermal/Vibration		20%

When a failure event occurs in HALT, a key distinction is made in identifying the event as either a hard failure, an operating limit failure or a destruct limit failure. This distinction guides the HALT process, indicating whether testing should proceed

at higher levels of stress or if sufficient progress has been made and the test should be stopped. Hard failures result from damage to the system's physical architecture requiring corrective action to restore performance to specification levels. Operating limit (OL) failures occur when one or more performance specifications aren't being met and cannot be brought back to within specification levels without reducing the stress level. It should be noted that the occurrence of an OL failure does not necessarily imply the occurrence of a hard failure. Destruct limit (DL) failures, also known as the fundamental limit of the technology (FLT), describe the occurrence of unavoidable failure modes due to the physical, chemical or structural limits of an item. Removing these failure modes typically requires a material substitution or significant structural change that is either physically impossible or financially infeasible. Examples of DL failures include the melting or phase change of a material, catastrophic mechanical failure due to vibrational stresses and dielectric breakdown due to voltage overstress. DL failures modes would not be considered representative of field experience and are useful only for driving system performance to the highest levels possible.

A product's operating margin is the stress range in which the system can operate for a specified mission duration without experiencing an operational limit failure. Figure 4 illustrates the relationship between a product's specification and its operating and destruct limits before and after a HALT. Comparing the top and bottom graphics in Figure 4 illustrates clearly the outward shift in the product's operating margin after the HALT. Additionally, Figure 4 suggests that the upper and

lower operating and destruct limits of a system are random quantities whose values can be represented by normal distributions with respective mean μ and variance σ. These parameters could then be estimated by the sample means $\hat{\mu}_i = \frac{\sum_{j=1}^{n} \max S_i}{n}$ and sample variances $\hat{\sigma}_i = \frac{\sum_{j}^{n} [\max S_i - \hat{\mu}]}{n-1}$ of the resulting maximum operating limit stress levels attained in testing each of the samples.

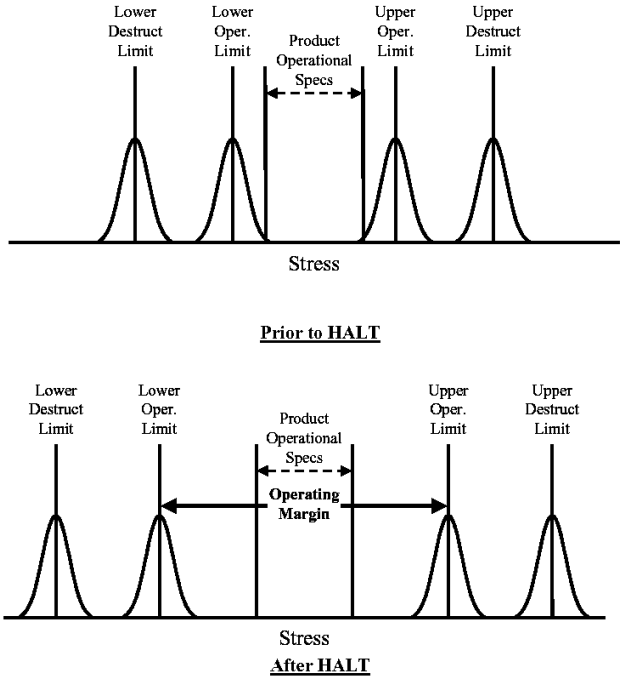

Figure 4 – Illustration of product specification, operating and destruct limits before and after highly accelerated life testing

Despite superficial similarities in terminology, HALT is disjoint from accelerated life testing (ALT) in several key areas. Whereas physical models have been developed [45] to compute acceleration factors that allow the inclusion of ALT test results into a reliability calculation, no such models exist for the inclusion of HALT test data due (1) to their inability to generalize across larger stress extrapolations, (2) the potential for complex stress profiles and (3) changes in the product's design

configuration. Nevertheless, practitioners have sought ways to predict field reliability from the results of HALT. Silverman [46] discussed the difficulties in analyzing HALT data using physics of failure models and acceleration factor calculations due primarily to limited data sharing across product categories allowing for more generalized approaches. Meeker [17] expressed concerns with the complexity of HALT stress profiles, advising practitioners to avoid analyzing HALT data using techniques derived from ALT. However, Meeker's discussion targeted tests in which multiple stresses were applied simultaneously and did not address simpler single stress test designs which are more common.

Sequential test procedures have been developed to utilize the results of HALT as a guide for follow-on tests to reduce the error involved in extrapolations across stress levels. Edson [47] introduced Calibrated Accelerated Life Testing (CALT) as a General Motors corporate standard to evaluate the reliability of products subject to HALT by adding test samples at key stress levels. The choice of the number of additional samples and stress levels to be tested in CALT are based on the maximum stress derived from the HALT. Bhote and Bhote [11] developed a nine-step life prediction modeling process known as multi-environment over-stress testing (MEOST). This method bypasses the independent stress tests and applies all of the stresses at once. Reliability estimation with MEOST is accomplished after observing a full year of service life to produce a useful estimate of the product's reliability

Brand and McLean were awarded three patents [48, 49, 50] comprising the proprietary HALT AFR Calculator® (AFR) model currently managed by the testing

firm Ops A La Carte®. The HALT AFR Calculator® is the most prominent HALT analysis model to-date utilizing linear, quadratic and exponential time acceleration techniques to estimate a product's relative life expressed as

$$R = \frac{MTBF_2}{MTBF_1} \tag{18}$$

where $MTBF_1$ is the field use reliability for the original unit subjected to an initial HALT and $MTBF_2$ is the field reliability for the unit subjected to a second HALT after implementing corrective action to remove the failure discovered during the first HALT. To utilize the HALT AFR Calculator it is assumed that 1) an estimate of the system's initial field reliability is known prior to testing, 2) that two rounds of HALT comprised of k independent stress tests have been performed and 3) the stress limits for both rounds of HALT are identical. Applying only the individual tests listed in Table 1, the AFR estimates relative life as a function the ratio of the times to first failure in HALT1 and HALT2 as

$$\hat{R} = \frac{f(t_{A1})}{f(t_{A2})} \tag{19}$$

where t_{A1} and t_{A2} are assumed to be realizations of independent exponential distributions and $f(\,)$ implies either a linear, quadratic or exponential function of the failure times, depending on which of the three patents is considered. Applying an appropriate acceleration factor (i.e. Arrhenius, Inverse Power Law, Eyring, etc.) gives the ratio of times to first failure in the un-accelerated time scale

$$\hat{R} = AF \cdot \frac{f(t_{A1})}{f(t_{A2})} = \frac{f(t_{F1})}{f(t_{F2})}. \tag{20}$$

Taking $\ln[\hat{R}]$ gives $\hat{R}^* = f(\mathrm{t}_{\mathrm{F}_1} - t_{F_2}) = f(\Delta t_F)$, for each of the k individual stressor tests, an independent $\hat{R}_k^*$ is obtained. The point estimate for the relative life R is then found from the simple average of these values and confidence limits are derived assuming that each $\hat{R}_k^*$ is approximately normally distributed.

Though McLean [51] has presented the results of the AFR Calculator to the accelerated test and reliability community, the manner in which the linear, quadratic and exponential acceleration techniques are utilized are unpublished.

Step Stress Accelerated Life Testing Data Analysis

Step-stress accelerated life testing (SSALT) involves testing items under successively higher levels of stress to drive failures quickly and reduce the overall testing time. The SSALT strategy avoids the problem inherent to constant stress accelerated life testing of selecting the stress levels to generate a sufficient number of failures to enable statistical inference. In SSALT, an initial load is applied at time $t = 0$, for a specified duration, after which the load is increased step-wise as time thresholds are reached. Testing stops when (1) a sufficient number of items on test have failed, (2) a limit stress has been reached or (3) the total time available for testing has been exhausted. The time duration that the units dwell at a given stress level may be based on the time required to conduct system diagnostics or on some optimality criterion. Bai [52] derived the optimal step time based on minimizing the asymptotic variance of the maximum-likelihood estimator of the mean life at a design stress. Laio [53] used the asymptotic variance of the estimated $100pth$ percentile of the product's lifetime distribution. In a HALT, equivalent time

increments are used and the length of each interval is based on the time required to conduct the necessary system performance diagnostics at each stress.

Estimating the lifetime of products at a use stress is substantially more difficult in products exposed to SSALT as the cumulative exposure at each stress level must be accounted for simultaneously. Two models commonly used for this purpose are the cumulative hazards model by Nelson [54] and the tampered failure rate model by Bhattacharyya and Soejoeti [19]. Some variations and generalizations of these models are presented by Kececioglu [55] and Zhao et al. [56]. In this literature review, only the Cumulative hazards model is discussed. It should be noted, however, that when the baseline failure distribution is exponential, the results obtained by both models coincide.

Cumulative Hazards Model

Consider the ordered sequence of increasing stresses s_i $(i = 1, 2, \dots)$ where the ratio s_{i+1}/s_i is constant with respect to time and the product's underlying life distribution at use stress s_0 is Weibull. If an accelerated stress s_1 is applied, the fraction of units failing up to time t is given by

$$F_1(t; s_1) = 1 - \exp\{-(Ks_1^n t)^\beta\}. \tag{21}$$

where the Weibull scale parameter $\alpha = Ks_1^{-n}$ reflects an inverse power law life-stress relationship with parameters K and n. If the stress level is changed to s_2 (21) is no longer appropriate to model system lifetime as it does not incorporate the cumulative exposure at both stress levels. Thus, to analyze the data from a step-stress test perspective, a cumulative exposure model is needed.

The cumulative hazards model [18] relates the life distribution of units under test at one stress level to the life distribution at another stress level. It is assumed that the remaining life of the test units depends only on the cumulative exposure the units have experienced up to the current time and is independent of how the exposure was accumulated. Moreover, since the units are held at a constant stress at each step, the surviving units will fail according to the distribution at the current step, but with a starting age corresponding to the total time accumulated prior to the beginning the current step. Thus, the probability of units failing during the time interval $(0, t_1)$ under stress s_1 not having experienced any other stresses may be described by (21). After time $t_1$the probability that the surviving units will fail during (t_1, t_2), while exposed to stress level s_2 is equivalent to the probability that the units would fail while at s_2 after accumulating $t - t_1$ plus an equivalent age ϵ_1 to account for the exposure at s_1 expressed as

$$F_2(t; s_2) = 1 - \exp\left[-\left(Ks_2^n\big((t - t_1) + \epsilon_1\big)\right)^{\beta}\right]. \tag{22}$$

The equivalent age ϵ_1 is the time at which the CDF for s_2 is equal to the CDF for s_1 after an exposure of t_1, or

$$\begin{aligned} F(t_1; s_1) &= F(\epsilon_1; s_2) \\ 1 - \exp\{-(Ks_1^n t_1)^{\beta}\} &= 1 - \exp\{-(Ks_2^n \epsilon_1)^{\beta}\} \\ s_1^n t_1 &= s_2^n \epsilon_1 \\ \epsilon_1 &= t_1\left(\frac{s_1}{s_2}\right)^n. \end{aligned} \tag{23}$$

This process may be used recursively to for an arbitrary number of stress levels where, in general the equivalent age is expressed as

$$\epsilon_i = t_i \left(\frac{s_i}{s_{i+1}} \right)^n + \epsilon_{i-1}. \tag{24}$$

Conclusion

This literature review has provided a framework for merging accelerated testing techniques with reliability growth methods by introducing several major components pertaining to the research that may have been unfamiliar to the reader. The Chapter has provided a synopsis of the research accomplished in the fields of reliability growth for complex systems, competing risks analysis and accelerated life testing data analysis. The parametric and nonparametric analysis of time to event data when competing risks are present was first discussed. Two statistical notions were presented as was the non-identifiability problems that can result when the risks cannot be assumed to be mutually independent. Reliability growth modeling was then discussed as were several concepts necessary to link traditional reliability growth modeling and accelerated testing to introduce accelerated reliability growth modeling. Finally, the cumulative damage model [18] was discussed as an approach to model time to failure in step-stress accelerated life testing. In the following Chapters, much of the information presented in this literature review is utilized to develop the reliability growth projection model to translate failure data obtained from a qualitative accelerated life test into to an estimate of system reliability after implementing corrective action.

III. Bridging the Gap between Quantitative and Qualitative Accelerated Life Tests

Jason K. Freels
Joseph J. Pignatiello
Richard L. Warr
Raymond R. Hill

Abstract – Test planners have long sought the ability to incorporate the results of highly accelerated life testing (HALT) into an early estimate of system reliability. While case-studies attest to the effectiveness of HALT in producing reliable products, the capability to translate the test's limited failure data into a meaningful measure of reliability improvement remains elusive. Further, a review of quality and reliability literature indicates that confusion exists over what defines a highly accelerated life test and how HALT differs from quantitative accelerated life testing (QALT) methods. Despite many authors making a clear distinction between qualitative and quantitative accelerated life tests, an explanation as to why this delineation exists cannot be found. In this paper, we consider an exemplary HALT composed of a single stressor to show that the HALT philosophy precludes the estimation of a system's hazard rate function parameters due to the test's fix implementation strategy. Four common accelerated failure data analysis methods are highlighted to show their limitations with respect to estimating reliability from HALT data. Finally, a way forward for future research is provided for improving the parameter estimation in follow-on testing.

Index Terms – competing risks, highly accelerated life testing, reliability growth, step-stress accelerated life testing

Introduction

Test planners have long sought the ability to incorporate the results of highly accelerated life testing (HALT) into an early estimate of system reliability. While case-studies [15, 16, 14] attest to the effectiveness of HALT in producing reliable products, the capability to translate the test's limited failure data into a meaningful measure of reliability improvement remains elusive. Further, a review of quality and reliability literature indicates that confusion exists over what defines a highly accelerated life test and how HALT differs from quantitative accelerated life test (QALT) methods. Despite authors making distinctions between qualitative and

quantitative accelerated life tests, a clear explanation as to why this delineation exists cannot be found in the literature.

The current paper introduces the purpose and process of a HALT before presenting an exemplary single-stressor highly accelerated life test to discuss the test's data structure. Then, approaches for characterizing product reliability are examined to highlight their limitations with respect to parameter estimation from HALT data. Finally, avenues for future research to utilize HALT data in follow-on testing are then suggested.

Highly Accelerated Life Testing

HALT originated in the electronics industry when Hobbs [44] began to use the term to describe a sequential testing process capable of quickly identifying the weak links in a product's design. More formally, HALT is an iterative test-fix-test [57] reliability growth process (Figure 5) wherein failure modes are discovered and removed as result of applying one or more stressors beyond the levels experienced in either the use environment or traditional accelerated life tests [12]. By subjecting products to such extreme environments many hard-to-find failure modes that would otherwise go undetected in traditional reliability tests, can be discovered and removed quickly. The HALT methodology is motivated by the tendency of nearly all products to experience one or more excursions outside of the expected usage environment over its lifetime. Such excursions can occur if the total stress exposure is not anticipated in product design (e.g. transportation stresses) or if the product is misused after fielding.

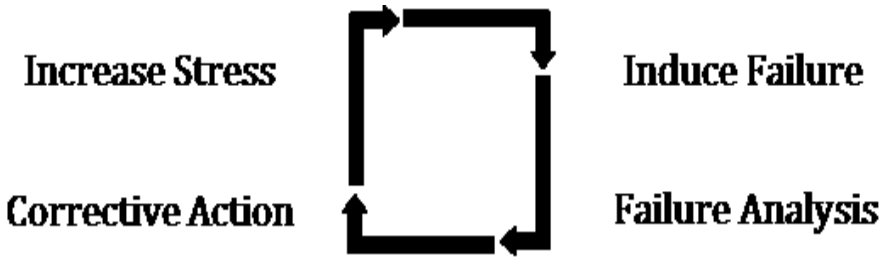

Figure 5 – Four-step highly accelerated life testing process

HALT improves a product's reliability by expanding its operating margin (Figure 6) as failure modes are discovered and removed, creating a "safety net" to withstand these excursions.

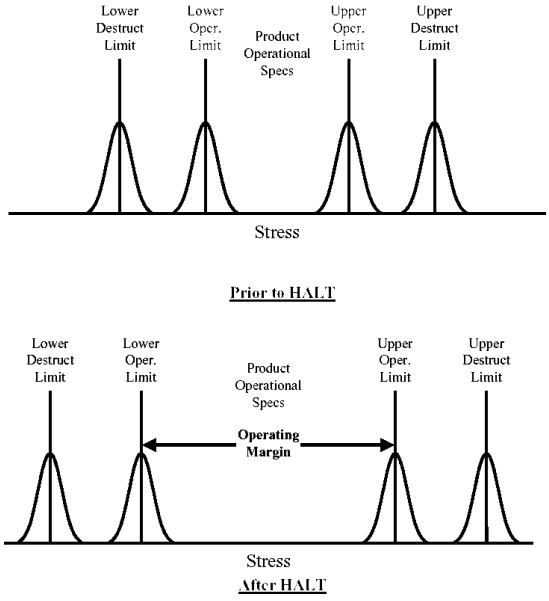

Figure 6 – Relationship between product operating and destruct limits

No formal specification exists to describe how HALT should be conducted, but a procedure described by Mclean [12] and used in industry is considered common practice [58, 59]. In this procedure, four to eight prototypes are exposed to an ordered stress regimen (Table 2). However, since the goal of HALT is to discover as

many latent failure modes as possible, test plans are neither limited to these stressors nor are they required to ensure that each stressor is included. Often product specific stresses (e.g. voltage/frequency margining) are added to or combined with these "standard" tests.

Table 2 – McLean HALT stress regimen

Stress Order	Stressor Type	Application Notes
1	Cold Step Stress	
2	Hot Step Stress	
3	Rapid Thermal Transitions	$\geq 60C^{o}$/minute rate of temperature change
4	Vibration Step Stresses	Simultaneous six degree of freedom random vibrations
5	Combined Rapid Thermal Transitions/Vibration	Vibration applied at thermal limits identified in Steps 1 and 2

During HALT, prototypes are electronically monitored to detect any degradation in performance or loss of function. When a failure occurs or a performance measure cannot be brought back within specification limits via stress reduction, corrective action is required to alter the system's architecture and remove the design weakness. As a result, the improved system can withstand higher stress levels prior to failure. This test-fix-test process continues iteratively as newer designs are stressed up to the physical or chemical limits of the unit (e.g. melting, dielectric breakdown). At this fundamental limit, corrective action requires a financially prohibitive design change, thereby concluding the test and finalizing the design. Upon release, the updated product will have reduced infant mortality and random

failure probabilities along with longer lifetimes with respect to wear-out in the designed usage environment as shown in Figure 7.

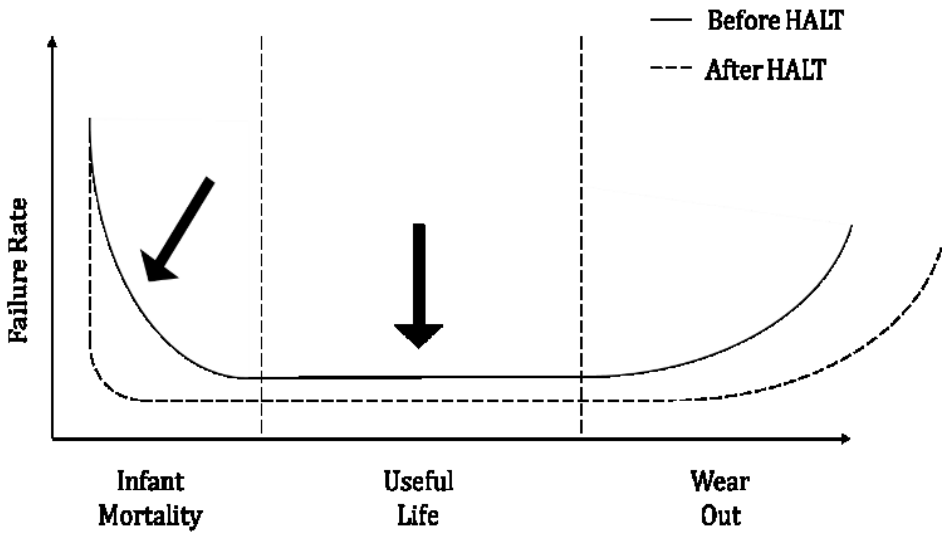

Figure 7 – Effect of highly accelerated life testing on bathtub curve

Example HALT Scenario and Data Structure

Meeker et al. [17] discuss the potential pitfalls of interpreting failure data produced from accelerated tests with complex stress profiles and multiple stressors. Here, an exemplary HALT composed of a single stressor is posited to illustrate that the HALT process, not the stress profile, precludes estimating a system's hazard rate function parameters. Consider a stress profile where HALT testing begins with the application of an initial stress level s_0 and increases step-wise (Figure 8) according to the piece-wise right-continuous function

$$s(t) = s_0 + \delta \cdot l(t) \tag{25}$$

where $l(t) = max\{i: \tau_i \leq t,\ i = 0, 1, \dots, M\}$, δ is the fixed step-up stress increment, and $[\tau_i, \tau_{i+1})$ represents the time interval in which the i^{th} stress level, s_i, is applied.

The sample set of units at risk in this exemplary HALT is comprised of m identical prototypes, sharing an initial design configuration $c_j = c_0$. Like most developed products, the prototypes are at risk of failure from multiple causes where each cause is itself the result of a flaw in the design activating in response to a given input. Some of these flaws may be found easily in that the range of stress inputs which are likely to expose them is large. However, many more flaws will be relatively difficult to find since only a small number of stress inputs will reveal their presence.

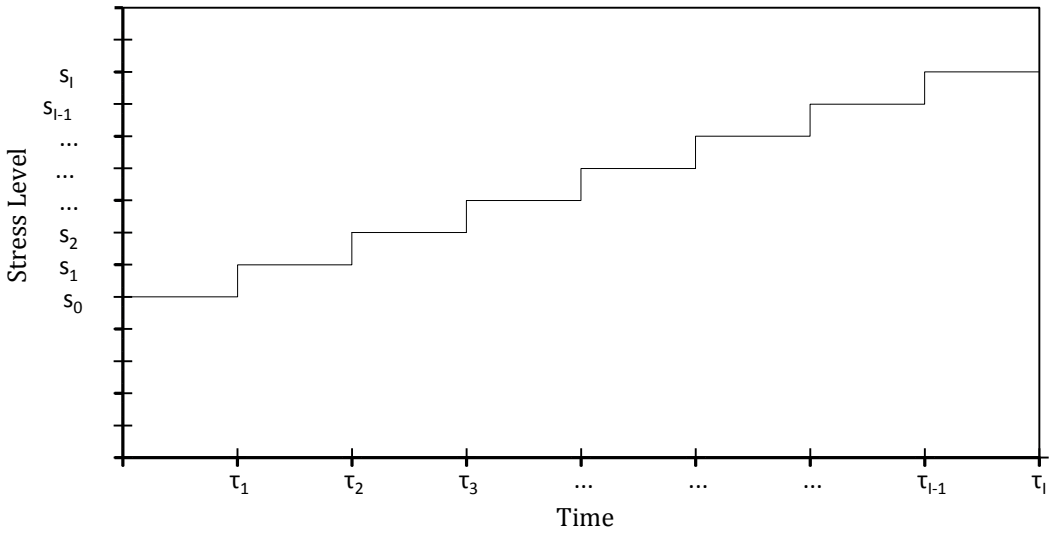

Figure 8 – Example step-stress profile

During $[\tau_i, \tau_{i+1}), i = 1, 2 \ldots, M$ an unknown number of latent failure modes, denoted by N, compete to be the cause of failure. Not all of the N overall modes are likely to be discovered as result of exposure to a single stressor (Figure 9), thus for the single-stressor HALT considered here we define $J \leq N$ as the number of failure modes ultimately discovered during the test.

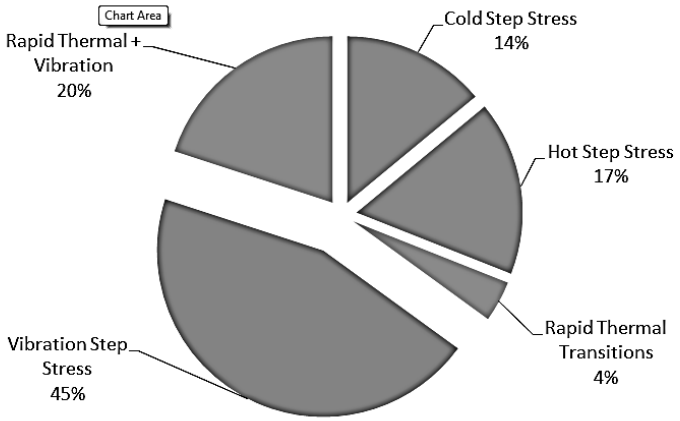

Figure 9 – Fraction of failures discovered by common HALT stressors (McLean, 2008)

This framework is akin to a serial arrangement of the failure modes where the occurrence of mode $j \in 1, \ldots, J$ at stress level s_i causes overall system failure and interrupts the test at clock time t_j where $\tau_i \leq t_j \leq \tau_{i+1}$. Implementing corrective action results in a new design configuration wherein failure mode j no longer exists, but additional failure modes not contained in the previous design, may have been created. Testing resumes on the updated design at stress s_i once the corrective action has been implemented on the unit under test as well as any remaining untested units. Consequently, the HALT test-fix-test strategy allows for a single observation of failure mode j prior to testing the new design and the data produced by each observed failure is comprised of the 4-tuple

$$\left[j,\ t_j,\ s_i(t_j), c_j\right] \tag{26}$$

representing the mode, time, stress level, and design configuration of the failure.

In the following section, four approaches cited in the literature for analyzing accelerated failure time data are discussed. The limitations of each approach are highlighted with regard to estimating the parameters of a product's reliability

function when the observed data is of the form shown in (26). The first and second subsections describe techniques commonly used for estimating the parameters of a product's reliability function where a specific failure mode of interest has been identified prior to testing. The third subsection describes an approach to extend traditional reliability growth models to incorporate data obtained at stress levels that are not representative of the end use environment. The fourth subsection discusses a commercial database used to predict system life by correlating the HALT results of multiple products with their subsequent field reliability. Finally, the fifth subsection suggests a direction for utilizing data of the form shown in (26) to improve system reliability estimation in follow-on testing.

Modeling Approaches

Quantitative Accelerated Life Testing Model

Quantitative accelerated life testing (QALT) [45] is widely used to expedite failures and quickly acquire information regarding a product's time to failure distribution. Test planners choose the type and levels of an accelerating variable along with the fraction of samples to be tested at each level to minimize the prediction variance of product life with respect to the use-level stress (V-optimality) [60]. The focus of QALT is typically on assessing the probability of failure due to specific wearout modes since a majority of infant mortality failures are removed prior to QALT through product ruggedization. For these wearout modes, knowledge of the relationship between the failure mechanism and the applied stress is required and may be available through an understanding of the chemical or physical

dynamics causing failure or through previous experience with similar systems. Using this prior information enables the selection of an appropriate life-stress model to quantify the manner in which the time to failure distribution changes with increased stress.

Analyzing the data of the form shown in (26) using QALT models requires coupling an appropriate underlying life distribution with a life-stress relationship [45] based on the physical/chemical dynamics causing failure [54]. Any distribution with a $[0, \infty)$ support region can serve as the underlying life distribution, however selection should be based on the distribution's ability to accurately describe the data as shown through goodness of fit testing [61]. For the example HALT described above, assume that the underlying life distribution relative to the use-level stress and failure mode j is $WEIB(\alpha_j, \beta_j)$ with CDF

$$F(t; \alpha_j, \beta_j) = P(T \leq t | \alpha_j, \beta_j) = 1 - e^{-\left(\frac{t}{\alpha_j}\right)^{\beta_j}} \quad (27)$$

where the value of the shape parameter β_j depends only on the failure mode j [62]. The true relationship between an accelerating variable and the failure mechanism can often be quite complicated, thus the selection of a life-stress relationship to model a failure process assumes a particular failure mode of interest has been identified prior to testing. Here, we assume an Inverse Power-Law stress-life relationship [63, 64] with parameters, $K, n > 0$ [55] reflected in the scale parameter α_j. Thus, under any constant stress s_i the fraction of units failing due to mode j prior to time t is

$$F_j(t; s_i) = 1 - \exp\left\{-\left(K_j s_i^{n_j} t\right)^{\beta_j}\right\}. \tag{28}$$

However, in a step-stress scenario (28) does not incorporate the exposure across multiple stress levels prior to failure and more complex techniques [19, 65, 54] are required. Nelson [18] introduced maximum likelihood estimation to step-stress QALT, suggesting a Weibull-Inverse Power Law likelihood function

$$L = \prod_i \left[\left\{ \beta_j K_j s_i^{n_j} \left(K_j s_i^{n_j} t\right)^{\beta_j - 1} e^{-\left(K_j s_i^{n_j}((t-\tau_i)+\epsilon_i)\right)^{\beta_j}} \right\}^{I_i} \times \left\{ e^{-\left(K_j s_i^{n_j}((t-\tau_i)+\epsilon_i)\right)^{\beta_j}} \right\}^{1-I_i} \right] \tag{29}$$

where I is an indicator function taking on a value of one if the observed value is a failure and zero it is a censored observation and ϵ_i incorporates the cumulative stress exposure prior to s_i. Relating failure times observed at elevated stresses to an equivalent use-stress exposure requires at least three observations of mode j before corrective action to estimate β_j, K_j and n_j. But this requirement violates the iterative test-fix-test nature of HALT which only allows for one observation for each mode prior to corrective action. Thus, without altering the HALT philosophy, the data shown in (26) is insufficient for estimating more than a single parameter using any life stress model or underlying life distribution.

Physics of Failure Analysis

Physics of Failure (PoF), or reliability physics, refers to a scientific approach wherein modeling and simulation are used to understand a product's reaction to external stressors and address its fitness for use with respect to the expected use conditions [66, 67]. The intent is to design-in reliability by investigating the

relationship between specific failure modes and accelerating variables and proactively eliminating the root causes of failure. The PoF approach provides the strongest characterization of system reliability [68], incorporating a diverse set of engineering disciplines along with physics, chemistry, metallurgy, mathematical statistics, and probability. PoF has gained wide acceptance among military and commercial sectors, showing significant time and cost [69] savings while allowing more information to be obtained during formal testing.

Pecht et al. [70] advocated for the use of PoF in reliability assessment in lieu of the popular parts count technique of Mil HDBK-217 [71]. Cushing et al. [72] identified several limitations of the parts count technique that could be addressed with PoF and presented procedures for implementing the approach (Table 3). Mendel [73] introduced probabilistic physics of failure (PPoF) as a technique to derive the statistical lifetime distribution and presented a case for applying PPoF in a design for reliability process. Modarres et al. [68] further emphasized that failure prediction is a probabilistic problem due to uncertainties associated with model parameters and failure-inducing agents that can result from changes in environmental, operating, and use conditions.

Alternatively, Snook et al. [74] identifies several limitations of PoF, most notably that for immature systems or those with multiple usage environments PoF analysis can be overly complex and burdensome. As case in point, Qi [75] presents a case study evaluating the fatigue life of printed circuit boards under thermal cycling.

Table 3 – Physics of failure process

Step	Process
1	Define realistic requirements
2	Define the expected mechanical, thermal, electrical and chemical usage environment experienced during manufacture, test, operation, storage, and repair
3	Identify potential failure mechanisms and the associated degradation processes (chemical, electrical, physical, mechanical, or thermal)
4	Define appropriate failure models and input parameters relative to the material characteristics, damage properties, relevant geometry, and operating environment
5	Compute the variability for each design parameter and the effective reliability function.
6	Design to the usage environment incorporating design stress spectra, part test spectra, and full-scale test spectra based on the anticipated life-cycle
7	Accept the design

Thirty thousand ball gate array (BGA) interconnections, arranged in 154 packages, were investigated to identify which were likely to cause the circuit boards to fall short of their 15-year lifetime requirement. To assess the likelihood of failure, three-dimensional finite element models, representing each distinct BGA package, were exposed to fifteen years of simulated field usage. The number of cycles to failure predicted by the Engelmaier-Wild solder creep-fatigue model [76] indicated that multiple BGA packages could not meet the fifteen year requirement.

Qi does not specify what corrective action was implemented subsequent to the investigation, but the case study illustrates that the level of effort required for PoF necessitates a mature design and a narrowly defined failure criteria. Physics of failure simulations are tailored to replicate specific product responses under a specific set of usage conditions and must be validated against demonstrated

performance. Had the circuit boards contained undiscovered failure modes, the value of the simulation data could be negated if corrective action alters the product's response to the applied stress. Identifying these latent modes through simulation involves replicating product responses under multiple stress environments where validation may be impossible with current physical models. Thus infant mortality failures must be removed to the fullest extent possible through a product ruggedization test such as HALT, prior to starting a PoF analysis.

Accelerated Reliability Growth

Reliability Growth is the positive improvement in a product's reliability over time due to design changes or manufacturing process improvements [9]. According to Ebeling [37], the earliest developed and most frequently used reliability growth model is that of Duane [38] who postulated from his empirical observations that for products under development, plots of cumulative failure rate versus cumulative operating hours on a log-log scale were approximately linear with negative slope. Crow [39] derived the stochastic basis for the Duane model, referred to in the literature as a power law process (PLP) or informally as a Weibull process – a type of nonhomogeneous Poisson process (NHPP) with a Weibull hazard intensity function $h(t) = \lambda\beta t^{\beta-1}$ where $\lambda > 0$ and $\beta > 0$ are interpreted as the initial reliability of the system and the growth rate parameter, respectively. The rigorous statistical development of the PLP allows estimation and prediction in both maximum likelihood and Bayesian contexts and has made the approach popular in the literature [6, 37].

Reliability growth testing has historically existed within the purview of large, complex development efforts, such as military weapon systems [77]. In these programs the stress profiles used in testing are assumed to reflect the field-use environment. Under these lower stress levels testing can extend several months and reliability growth planning curves (Figure 10) are used to ensure adequate test-time is made available to discover failures and subsequently attain the reliability requirement with some level of confidence. Testing exposes several distinct failure modes, thus traditional reliability growth models track only the overall number of failure events occurring in a test phase. Reliability improvement is therefore shown as a decrease in the rate of failure events across test phases (assessed growth). An indication that the reliability growth is lagging (assessments falling below the planning curve) may signal the need to reallocate test resources or reevaluate the testing strategy.

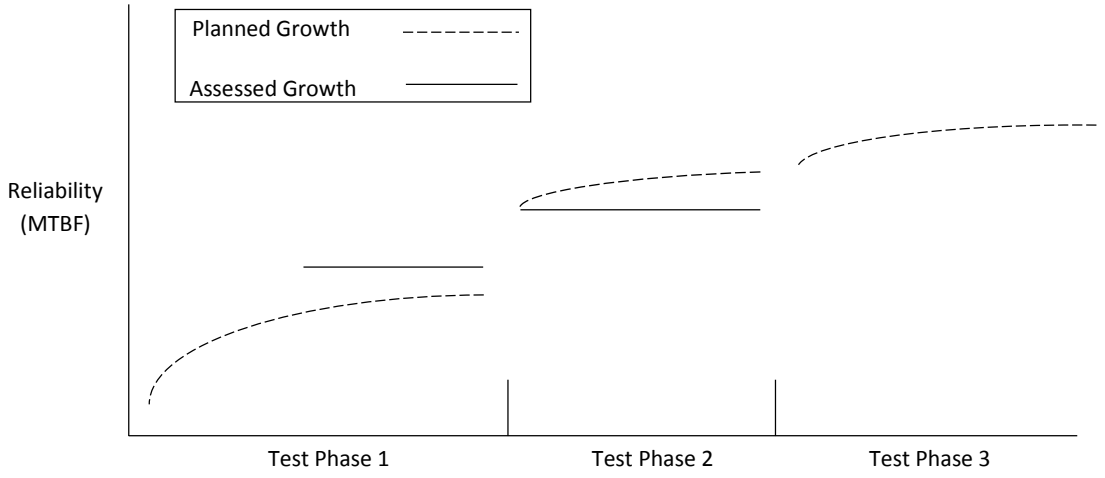

Figure 10 – Idealized reliability growth planning curves developed across three test phases showing the difference in planned and assessed reliability (Source: MIL-HDBK-189C, p. 11)

In commercial environments several products types may be under development simultaneously, thus little time may exist to ruggedize and quantify the reliability of a particular product. In these situations, accelerated testing methods such as HALT may be utilized as part of the reliability improvement process prior to market introduction. Like traditional growth testing, several distinct failure modes are observed, but many of these observations will only be realized at elevated stress levels. The Army Materiel Sustainment and Analysis Activity (AMSAA) [9] notes that traditional reliability growth models produce inaccurate results when the stress profile is not operationally relevant. Strunz and Herrmann [78] address planning, tracking and projecting reliability growth of liquid rocket engines where the test profile does not reflect the operational use environment. Concluding that the traditional modeling approaches are insufficient, the authors derive a Bayesian estimation method that accounts for the characteristics of the test profile and aggregates component, subsystem and system data.

Feinberg [77] introduced accelerated reliability growth testing (ARGT) by modifying the AMSAA [9] reliability growth planning model equations to include a system-level acceleration factor. The approach is based on the assumption of a linear correspondence between the reliability growth attained under an accelerated stress and that which would have occurred using only the use-level stress. Feinberg does not qualify this assumption, implying he expects it to hold regardless of the intensity of the accelerated stress and the level of extrapolation considered. Intuitively, as stress increases the time to failure for a given mode decreases.

However, as noted by [45] the time compression is not equivalent for each failure mode in the system. While Feinberg indicates that acceleration factor models are mode specific, he assumes a system-level acceleration factor can be derived, but leaves this task to the reader.

With respect to estimating product reliability from the data in (2), reliability growth models improve upon QALT and PoF in that design changes can be accommodated. However, traditional models assume growth *a priori* via a fix effectiveness value applied across each corrective action. In an accelerated test, this assumption is invalid as failures may be discovered far from the use stress. Clearly, the reliability growth attained through corrective action depends on the likelihood of the failure actually occurring in operation [54, 79, 17]. Nelson [54] discusses this phenomenon, describing a costly effort to remove a failure mode which never would have occurred in actual use, resulting in wasted resources and no actual reliability improvement. Moreover, relating growth to a decrease in failure events is also invalid as the rate of failures may rise with increased stress even though reliability is improving.

McLean's AFR Estimator Model

After collecting HALT data from more than fifty products across twenty industries McLean [80] developed the *Actual field Failure Rate (AFR) Estimator* to correlate a product's HALT results with its subsequent field failure history. When furnished with the appropriate HALT data and product information (Table 4) the

model provides the field failure rate for the improved design with confidence limits based on χ^2 estimates derived from the SEMI E10 standard [81].

Table 4 – AFR Estimator Inputs

A	Initial MTBF estimate (early testing, reliability prediction standards)
B	HALT Hot operating limit as measured on the product
C	HALT Cold operating limit as measured on the product
D	HALT Vibration operating limit as measured on the product
E	Published thermal operating specifications (in degrees Celsius)
F	Sample size used in the final HALT
G	Field duty cycle

The AFR Estimator utilizes linear, quadratic and exponential time acceleration techniques to estimate a product's relative life,

$$R = \frac{MTBF_{redesign}}{MTBF_{original}} \tag{30}$$

representing the proportional improvement in the mean time between failures of the redesigned unit over that of the original design. The model also provides recommended minimum HALT stress levels to ensure sufficient test coverage in each environment and to assure the product will exceed customer expectations. Notwithstanding these advantages, the failure rate predictions can only be made on the basis of the temperature and vibration tests. Further, the model can't estimate wear-out mechanisms which must be addressed using QALT techniques. Although McLean has presented results [51] of the AFR Estimator, the model remains unpublished and is proprietary to Ops A La Carte® LLC. Thus, the mechanics of how

the linear, quadratic and exponential acceleration techniques are utilized is unclear. Interestingly, Brand and McLean own three patents [48, 49, 50] outlining models for reliability estimation after a HALT using linear, quadratic and exponential acceleration techniques. However, correspondence with the author indicates that the AFR Estimator is distinct from the models described in these patents.

Future Research Avenues

Current accelerated reliability techniques investigate the response of a known failure mode in the presence of one or more accelerating variables assuming the system is mature. Estimating reliability from data obtained from qualitative accelerated life tests [11, 42, 13] depends on identifying a relationship to interpret the observations made on early configurations to those of the final design. This relationship should naturally be a function of the original degradation process, but must also incorporate changes resulting from corrective actions. However, for even a single failure mode-stressor combination, this relationship can be extremely complex as each corrective action can result in a new degradation process with unfamiliar failure states. Although reliability physics models continue to improve, the push for testing chambers that can produce harsher environments, find failures more quickly, and test more complex products outpaces the capacity to derive tools that can accurately describe product behavior in these environments. Consequently, HALT data is difficult to obtain as few entities have the equipment necessary to conduct the test and the failure information is often considered competition sensitive to both the item's manufacturer and the testing organization.

In the absence of sufficient physical models and available data, progress in this area will likely be confined to scanning the multi-dimensional stress space bound by the HALT destruct limits or developing novel competing risk models based on reasonable, but unverifiable, assumptions on the form of the data and model propriety at the highest stresses. Escobar and Meeker [60] investigated optimal accelerated test designs when two or more independent accelerating variables are present. Gao et al. [82] compared the Escobar-Meeker approach to orthogonal and uniform designs in computer experiments under V-optimality. Their results showed that the Escobar-Meeker approach performed best in estimating the parameters and the p^{th} quantile of the life distribution under normal stress.

Assuming independent experimental factors may be reasonable in the neighborhood of the design stress. But at limit stress failure depends on the total energy put into the system. In this circumstance the design space is non-quadrangular with an outer surface representing the zero-sum relationship between the factors. Although this relationship is unknown for many combinations of experimental factors, comparing the Escobar–Meeker, orthogonal and uniform designs under arbitrary design spaces and dependence relationships would be a valuable contribution. Additionally, extending methods such as Calibrated Accelerated Life Testing (CALT) [47] to multiple dimensions may prove beneficial.

Conclusion

Maintaining market share compels manufacturers to provide evolutionary and revolutionary capabilities under aggressive development schedules while ensuring

a high reliability in complex stress environments. Successfully attaining reliability requirements for these products requires setting and achieving reliability targets throughout the various stages of the development process. However, schedule and cost constraints often preclude sufficient testing on early prototypes to generate meaningful reliability estimates. Under these constraints, incorporating data obtained throughout a development effort can result in improved reliability estimation prior to fielding. Qualitative accelerated tests are often used to quickly improve system reliability by identifying and removing initial design flaws. Generally, no attempt is made to produce a reliability measure from the limited data obtained in these qualitative tests as relevant approaches require more data. The delineation between qualitative accelerated life tests and QALT exists due to the philosophy of removing the root cause of each failure through corrective action. However, the output of qualitative testing can be used to improve reliability distribution parameter estimation and explore the outer surface of the enlarged design space to guide follow-on tests.

IV. Maximum Likelihood Estimation for the Poly-Weibull Distribution

Jason K. Freels
Joseph J. Pignatiello
Richard L. Warr
Raymond R. Hill

ABSTRACT

The Weibull distribution has long been a popular choice for modeling lifetime data of various mechanical and biological phenomena when the associated hazard rate function is constant or monotone increasing or decreasing. However, non-monotone hazard functions are common in reliability and survivability contexts where a system may undergo an initial "burn-in" prior to periods of useful life and eventual wearout. In these scenarios, the standard Weibull can only model a portion of the "bathtub" curve but is incapable of adequately modeling the entire failure process. Several modifications to the standard two-parameter Weibull distribution have therefore been introduced in the literature to effectively model and analyze lifetime data where the hazard rate function is bathtub shaped. The performance of each modified distribution is typically assessed by its ability to fit reference data sets that are known to have a bathtub shaped hazard rate function. The current paper compares the performance of two recent contributions in this area to that of the poly-Weibull distribution with respect to several goodness of fit measures. In addition, numerical and analytical procedures are developed for obtaining the maximum likelihood parameter estimates and standard errors for the generalized poly-Weibull distribution with arbitrary number of terms. Our results show that both the bi-Weibull and tri-Weibull distributions fit the reference data sets better that either of the current best-fit models.

Introduction

The Weibull distribution has long been a popular choice for modeling lifetime data of various mechanical and biological phenomena when the associated hazard rate function is either constant or monotone increasing or decreasing. However, non-monotone hazard functions are common in reliability and survivability contexts where a system may undergo an initial "burn-in" prior to periods of useful life and eventual wearout. In these scenarios, the standard Weibull distribution can model a portion of the "bathtub" hazard curve but is incapable of describing the entire

failure process. Several modifications to the standard two-parameter Weibull distribution have therefore been introduced in the literature to effectively model lifetime data where the hazard rate function is bathtub shaped. The performance of each modification is assessed by comparing its goodness of fit to data sets published by Aarset [83] and Meeker [84], known to have bathtub shaped hazard functions.

Bathtub-shaped hazard functions arise from the existence of multiple competing failure modes which dominate at different epochs in the life of a system. The poly-Weibull [85] distribution arises naturally in scenarios of competing risks as it describes the minimum of several independent random variables where each follows a distinct Weibull law. The current paper compares the goodness of fit of the poly-Weibull to two recently proposed Weibull modifications, the new modified Weibull distribution [86] and the exponentiated modified Weibull extension distribution [87]. Our results show that the poly-Weibull fits both data sets better than either of these modifications.

Almalki and Yuan [86] introduced the new modified Weibull (NMW) distribution

$$F_{NMW}(t;\alpha,\theta,\beta,\gamma,\lambda) = 1 - e^{-\alpha t^{\theta} - \beta t^{\gamma} e^{\lambda t}} \quad \alpha,\theta,\beta,\gamma,\lambda \geq 0, \quad t \geq 0 \tag{31}$$

by considering a two-component serial arrangement in which one component follows a standard two-parameter Weibull model and the other follows a modified Weibull (MW) distribution [88]. The authors show that the goodness of fit of the NMW is superior to that of other four and five parameter Weibull modifications including the additive Weibull distribution [89], the modified Weibull distribution

[90] and the beta modified Weibull distribution [91]. Likewise, Sarhan and Apaloo [87] showed that the exponentiated modified Weibull extension (EWME) model

$$F_{EWME}(t;\lambda,\alpha,\beta,\gamma) = \left[1 - e^{\lambda\alpha\left(1-e^{(t/\alpha)^{\beta}}\right)}\right]^{\gamma} \quad \lambda,\alpha,\beta,\gamma \geq 0, \quad t \geq 0 \tag{32}$$

fit the reference data better than the exponentiated Weibull [92], the exponentiated Gompertz [93] and the modified Weibull extension distributions [94].

The Poly-Weibull Distribution

The CDF of the poly-Weibull distribution is expressed as:

$$F_{PW}(t;\boldsymbol{\alpha},\boldsymbol{\beta}) = 1 - \left\{\exp\left[-\sum_{j=1}^{J}\left(\frac{t}{\alpha_j}\right)^{\beta_j}\right]\right\} \quad \alpha_j,\beta_j > 0, \quad t \geq 0 \tag{33}$$

where α_j, β_j represent the scale and shape parameters associated with the Weibull model describing risk $j = 1, \ldots, J$. Accordingly, the poly-Weibull density function is

$$f_{PW}(t|\boldsymbol{\alpha},\boldsymbol{\beta}) = R(t)h(t) = \left\{\exp\left[-\sum_{j=1}^{J}\left(\frac{t}{\alpha_j}\right)^{\beta_j}\right]\right\} * \sum_{j=1}^{J}\frac{\beta_j t^{\beta_{j-1}}}{\alpha_j^{\beta_j}}. \tag{34}$$

When $J = 2$, equations (33) and (34) are the CDF and pdf of the bi-Weibull distribution, and when $J = 3$ the model is naturally known as the tri-Weibull distribution. Considering the poly-Weibull was introduced over twenty years ago [85], the body of literature studying its properties is quite limited. Within this body, the predominant focus has been on Bayesian estimation of the bi-Weibull model parameters. Berger and Sun [85] studied the bi-Weibull distribution using a Gibbs sampling algorithm with informative and non-informative priors. The authors computed the posterior density and predictive reliability when the shape parameters are either known or unknown. Their simulation results compared

favorably with a closed-form exact solution, but were shown to be computationally expensive for small sample sizes. Davison and Louzada-Neto [95] also explored the bi-Weibull model, however, they assert that Markov chain simulation is not necessary for obtaining posterior probabilities. Using real and generated data, the authors illustrate, that either Laplace's method [96] or the Bayesian bootstrap [97] are sufficient and can reduce the computational burden.

Figures 11 and 12 illustrate several bathtub shapes modeled with the bi-Weibull and tri-Weibull distributions with parameter vector $\theta = [\beta_1, \ldots, \beta_J, \alpha_1, \ldots, \alpha_J]$.

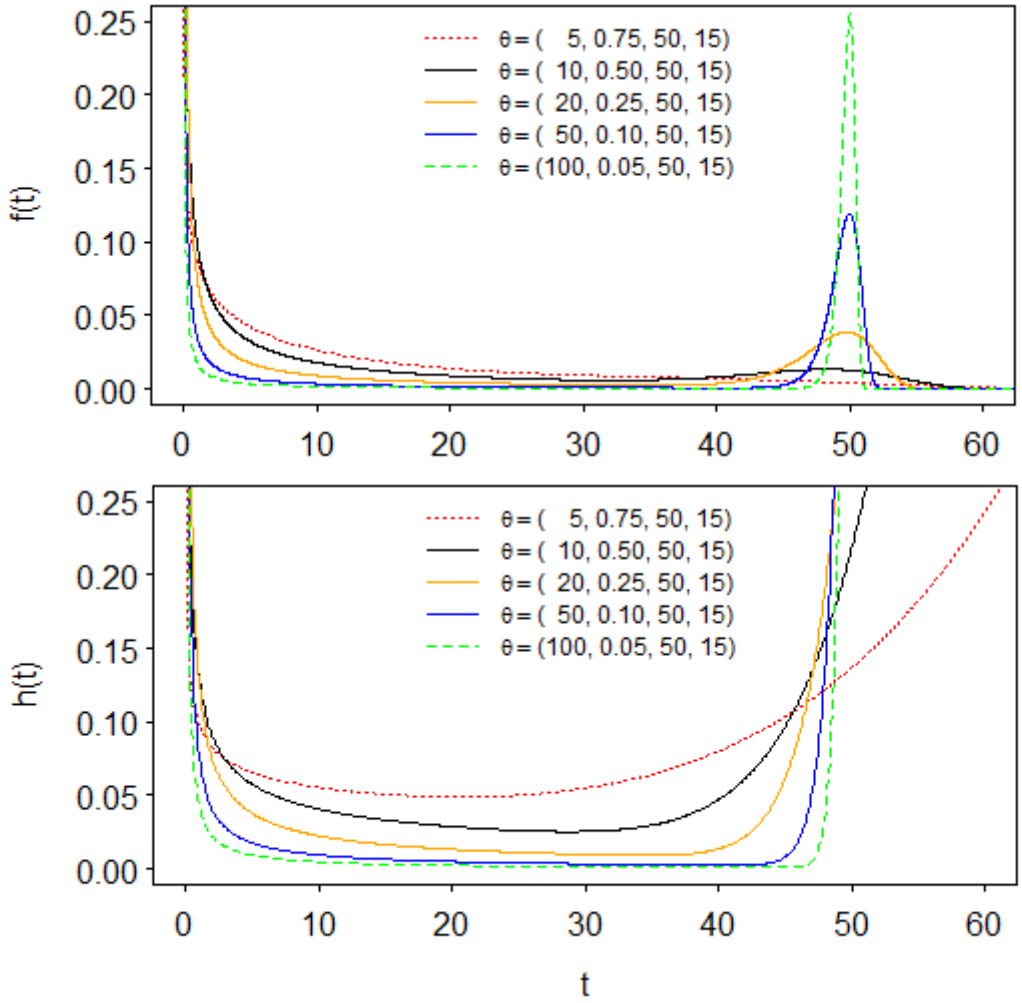

Figure 11 – Example bi-Weibull density (top) and hazard (bottom) functions

Comparing the center sections of each figure, the reader will note the ability of the tri-Weibull to model complex failure processes for which the "useful life" portion of the hazard function is not constant and thus does not follow an exponential model.

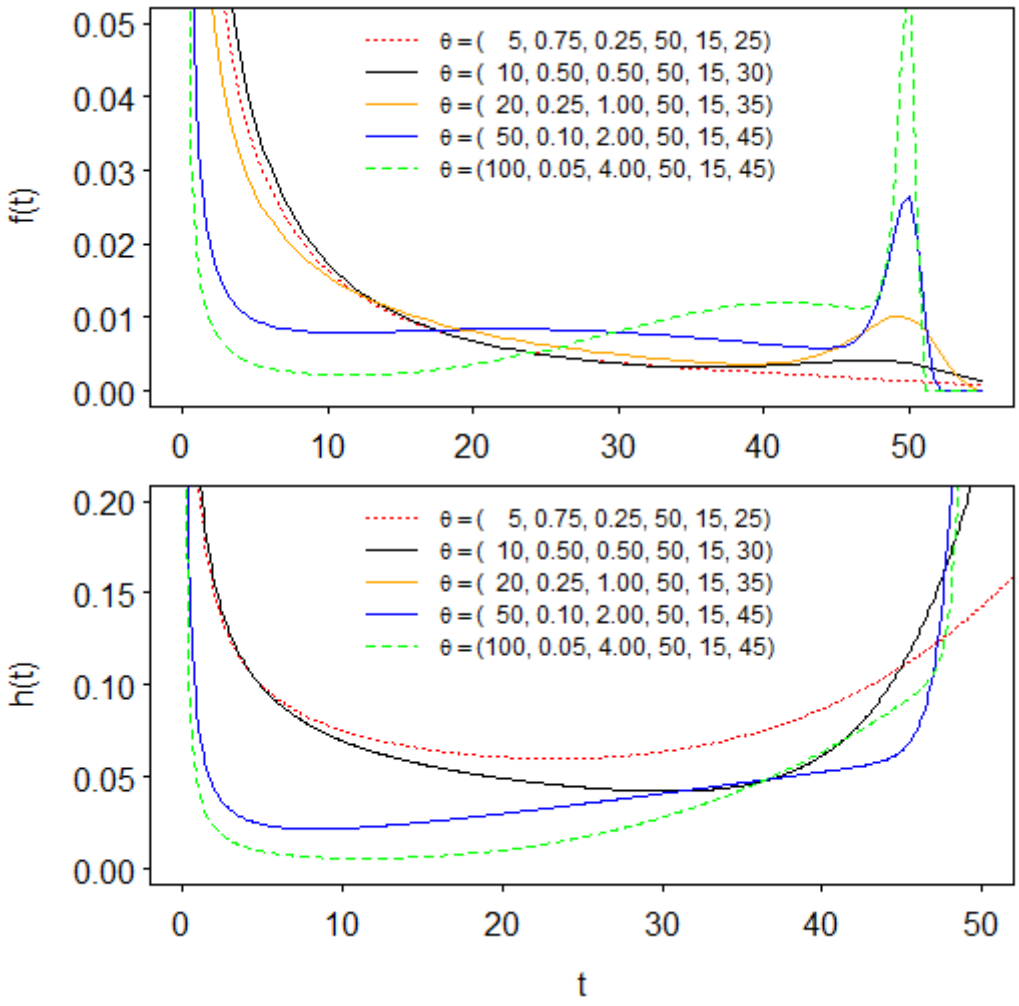

Figure 12 – Example tri-Weibull density (top) and hazard (bottom) functions

An advantage of the poly-Weibull model is that the parameter values have the same relationship to the mean and variance of the failure process as does the standard two-parameter Weibull model. Thus $\beta_j < 1$ implies a failure process with a large coefficient of variation and decreasing hazard rate indicative of infant mortality.

Conversely, $\beta_j > 1$ portends a wear-out failure mechanism with an increasing hazard rate function reflective of a lower coefficient of variation. The poly-Weibull distribution is therefore capable of generating bathtub shaped hazard functions by modeling multiple failure processes simultaneously as Figures 11 and 12 illustrate.

Parameter Estimation

Consider a random sample of observations $t_1, \dots, t_n$ from a poly-Weibull($\boldsymbol{\alpha}, \boldsymbol{\beta}$) distribution with unknown parameter vector $\theta = (\alpha_1, \dots, \alpha_J, \beta_1, \dots, \beta_J)$, and indicator variable δ_i where $\delta_i = 1$ if t_i is a failure time and $\delta_i = 0$ if t_i is a censoring time. The log-likelihood function $\mathcal{L}(\theta)$ based on data $(t_1, \delta_1), \dots, (t_n, \delta_n)$ is

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \left[\delta_i \log \sum_{j=1}^{J} \left(\beta_j t^{\beta_j - 1} \alpha_j^{-\beta_j} \right) - \sum_{j=1}^{J} \left(\frac{t_i}{\theta_j} \right)^{\beta_j} \right]. \tag{35}$$

Davison and Louzada-Neto [95] initially presented the system of $2J$ nonlinear equations, which when solved would provide the maximum likelihood estimates for α_j and β_j, $j = 1, \dots, J$. However, these equations contained an error – omitting β_j from the denominator of $\partial\mathcal{L}/\partial\alpha_j$. The corrected equations for obtaining the poly-Weibull MLE's, which have been verified both analytically and numerically, are therefore

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha_j} &= \sum_{i=1}^{M} \left(\frac{t_i}{\beta_j} - \frac{\delta_i}{h(t_i)} \right) \left(\frac{\beta_j}{\alpha_j} \right) h_j(t_i) = 0 \qquad j = 1, \dots, J \\ \frac{\partial \mathcal{L}}{\partial \beta_j} &= \sum_{i}^{M} \left[\delta_i \frac{h_j(t_i)}{h(t_i)} \left\{ \beta_j^{-1} + \log\left(\frac{t_i}{\alpha_j} \right) \right\} - \frac{t_i}{\beta_j} h_j(t_i) \log\left(\frac{t_i}{\alpha_j} \right) \right] = 0 \qquad j = 1, \dots, J \end{aligned} \tag{36}$$

where $h_j(t_i) = \beta_j t^{\beta_j - 1} \alpha_j^{-\beta_j}$ and $h(t_i) = \sum_{j=1}^{J} \left(\beta_j t^{\beta_j - 1} \alpha_j^{-\beta_j} \right)$. Solving this system of non-linear equations cannot be accomplished analytically and the use of Newtonian or quasi-Newtonian numerical optimization techniques can be tedious as finding a solution is highly sensitive on the starting values for the parameters in each equation. We find it simpler to obtain accurate parameter estimates by maximizing the log-likelihood function directly using a quasi-Newtonian algorithm [98]. However, for asymptotic interval estimation, the optimization algorithm can produce inaccurate Hessian matrices leading to negative values along the diagonal of the covariance matrix. Thus for finding the standard errors of the poly-Weibull model parameters the components of the observed Fisher information matrix $\mathcal{I}(\theta^*) = -\nabla\nabla^{\mathrm{T}} \mathcal{L}(\theta)|_{\theta=\theta^*}$ have been derived analytically and are presented in appendix A.

Application

In this section the Aarset [83] and Meeker [84] data sets are analyzed and the goodness of fits for the bi-Weibull and tri-Weibull distributions are compared to those of the NMW [86] and the EMWE [87]. In keeping with the established precedent, the goodness of fit comparison in this paper is based on the values of the Kolmogorov-Smirnov (K-S) test statistic, the Akaike information criterion (AIC) and the log-likelihood function for each model computed at their respective MLE's. In addition, the fit of each distribution is assessed graphically as the reliability, density and hazard functions are plotted against their nonparametric counterparts.

Aarset data

The Aarset [83] data set (Table 5) represents the lifetimes of 50 devices and contains no censored observations.

Table 5 – The Aarset data set

0.1	0.2	1	1	1	1	1	2	3	6
7	11	12	18	18	18	18	18	21	32
36	40	45	46	47	50	55	60	63	63
67	67	67	67	72	75	79	82	82	83
84	84	84	85	85	85	85	85	86	86

Table 6 shows the MLE's and standard errors for the parameters of each of the four models considered while Table 7 displays a comparison of each model's goodness of fit measures. The data in Table 7 indicates that the null hypothesis of the two-sample K-S test cannot be rejected for any of the four models at a significance level below 0.8. However, the data also shows that the tri-Weibull and bi-Weibull fit the data better than either the NMW or the EMWE as both have larger likelihoods as well as smaller K-S statistics and AIC values. Further, the superior performances of the poly-Weibull models are immediately clear upon observing Figure 13. In the top plot, the reliability function of each model is plotted against the Kaplan-Meier [30] nonparametric estimate of the reliability function for the data.

Table 6 – Maximum likelihood estimates (standard errors) for the Aarset data

Model	MLE of the Parameters					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
Tri-Weibull	98.152	0.524	4.215	85.091	122.478	92.299
	(32.762)	(0.056)	(0.937)	(0.339)	(52.740)	(8.870)
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$		
Bi-Weibull	82.334	0.702	84.907	61.663		
	(22.602)	(0.075)	(0.328)	(14.538)		
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\gamma}$		
EMWE	49.050	7.18×10^{-5}	3.148	0.145		
	*	*	*	*		
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	
NMW	0.071	0.197	7.015×10^{-8}	0.016	0.595	
	(0.031)	(0.184)	(1.501×10^{-7})	(3.602)	(0.128)	

The NMW and EWME fit the central portion of the data better that the bi-Weibull while both the bi-Weibull and tri-Weibull more closely fit the upper and lower tails of the data where the majority of observations are concentrated. However, the tri-Weibull is clearly the best fit throughout the entire range of the observations.

Table 7 – Performance measures for the Aarset data set

Model	Params	Log-Lik	K-S	p-value	AIC
Tri-Weibull	6	-202.51	0.063	0.998	417.01
Bi-Weibull	4	-206.09	0.100	0.925	420.20
EMWE	4	-213.86	0.101	0.646	435.72
NMW	5	-212.90	0.088	0.803	435.80

In the middle plot, each model's density function is plotted against a histogram of the data. Note that the poly-Weibull models indicate that the probability of failure after the final observation is near zero, as would be expected for a system with a true bathtub–shaped hazard, while the NMW and EMWE do not reflect this. This is also clearly evident in the bottom plot where the hazard functions are plotted against the empirical hazard plot [99].

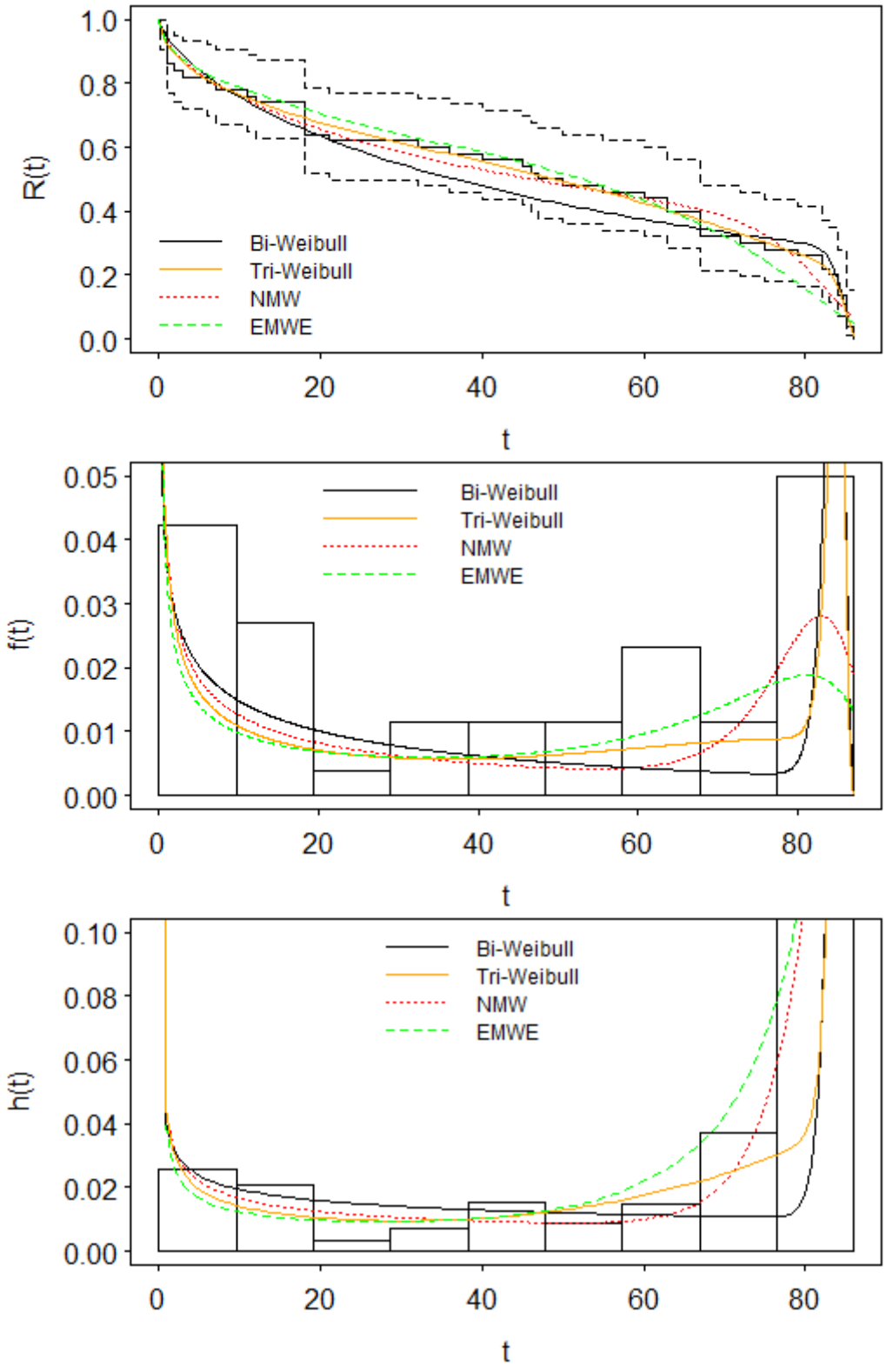

Figure 13 – Triple plot illustrating the fit of the bi-Weibull, tri-Weibull NMW and EMWE models to the Aarset data

Meeker data - Uncensored

The Meeker [84] data set (Table 8) represents the observed lifetimes of 30 devices and includes eight censored observations. The data set only has a bathtub shaped hazard function if the censored observations are treated as failures. Thus, the subsequent analysis follows this precedent established in the literature. In the following subsection the Meeker data is analyzed again, correctly treating the final observations as censored.

Table 8 – The Meeker data set

2	10	13	23	23	28	30	65	80	88
106	143	147	173	181	212	245	247	261	266
275	293	300+	300+	300+	300+	300+	300+	300+	300+

Maximum likelihood parameter estimates and standard errors are displayed in Table 9. Note the large standard errors associated with the bi-Weibull and tri-Weibull shape parameter $\beta_{j=1}$. As a result of treating the censored observations in the Meeker data set as failures we have enforced a deterministic failure at 300 hours such that

$$P(T > 300|T = 300) = 0. \tag{37}$$

Thus, the optimization procedure attempts to fit bi-Weibull and tri-Weibull distributions to the uncensored Meeker data with shape parameter $\beta_1 \to \infty$. The large standard errors therefore reflect the difference between the true distribution governing the failure process and the distribution that can be numerically obtained. Observing the density curves in the center plot of Figure 14, shows that the poly-

Weibull accurately portrays a deterministic failure at 300 hours while the NMW and EWME curves incorrectly show that failures beyond 300 hours are possible.

Table 9 – Maximum likelihood estimates (standard errors) for the uncensored Meeker data set

Model	MLE of the Parameters					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
Tri-Weibull	124.000	5.591	0.738	299.850	356.022	352.185
	(188.557)	(1.907)	(0.102)	(1.825)	(46.686)	(136.733)
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$		
Bi-Weibull	124.000	0.892	299.780	253.940		
	(98.228)	(0.122)	(1.117)	(62.099)		
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\gamma}$		
EMWE	197.210	5.468 x 10^{-6}	4.482	0.129		
	*	*	*	*		
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	
NMW	0.024	0.056	5.991 x 10^{-8}	0.012	0.629	
	(0.019)	(0.024)	(8.164 x 10^{-8})	(1.290)	(0.158)	

Regardless, the goodness of fit measures in Table 10 indicate that bi- and tri-Weibull remain the preferred models when compared to the NMW and EMWE for fitting the data. With additional computing power this fit can be improved.

Table 10 – Performance measures for the uncensored Meeker data set

Model	Params	Log-Lik	K-S	p-value	AIC
Bi-Weibull	4	-156.14	0.105	1.000	320.30
Tri- Weibull	6	-154.65	0.097	1.000	321.30
EMWE	4	-166.35	0.131	0.632	340.71
NMW	5	-166.18	0.148	0.482	344.40

The graphical fit, as shown in Figure 14, again indicates that both the tri-Weibull and the bi-Weibull outperform the NMW and EMWE. However, because the bathtub shape is less marked for the uncensored Meeker data, the difference between the bi- and tri-Weibull is insufficient to overcome the penalty of two additional parameters.

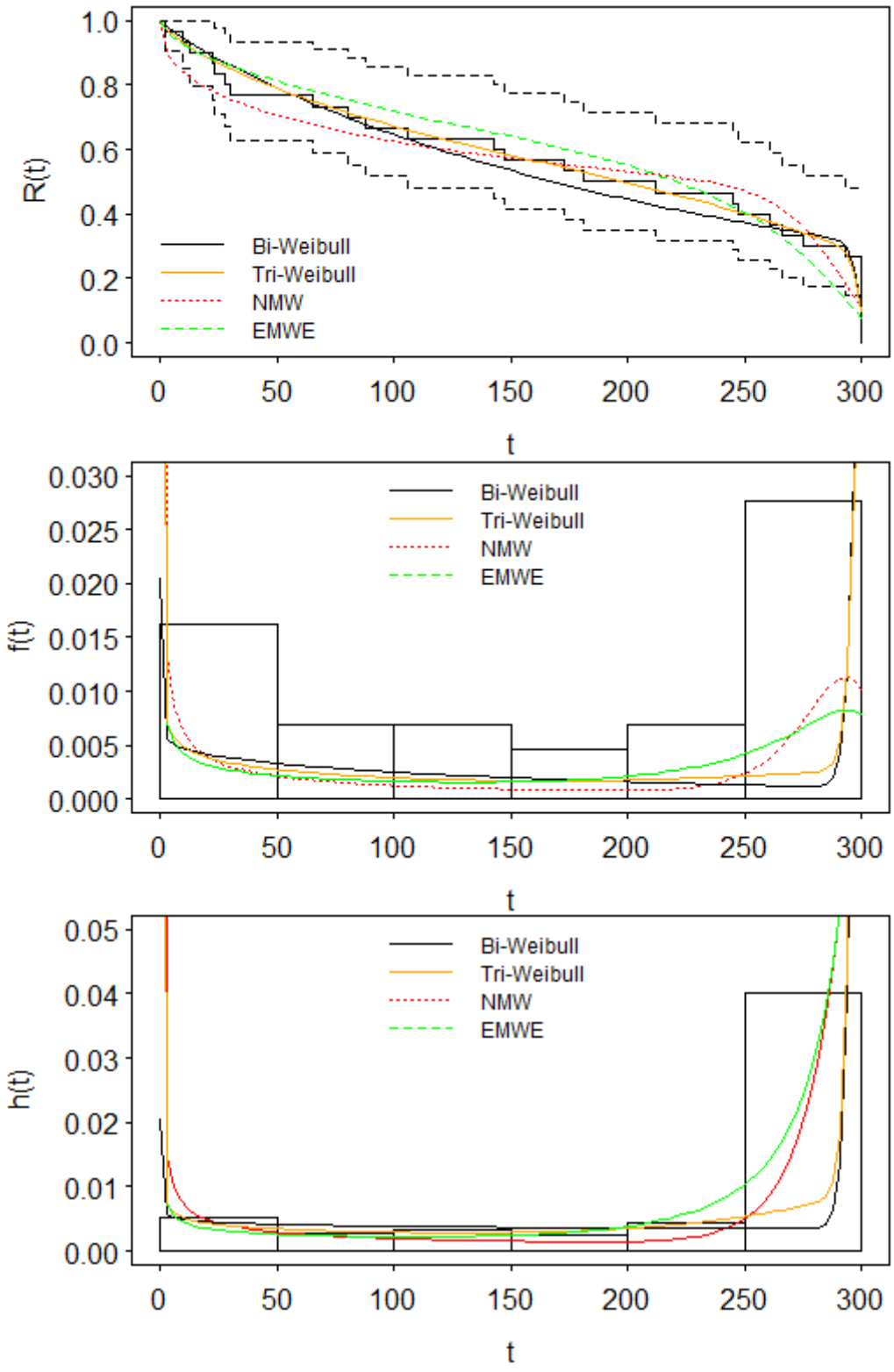

Figure 14 – Triple plot illustrating the fit of the bi-Weibull, tri-Weibull NMW and EMWE models to the uncensored Meeker data

Meeker data - Censored

The Meeker [84] data set is analyzed again, this time treating the final eight observations as right censored. As the respective authors do not present results for the NMW and EMWE in this scenario, the results presented below were obtained using the numerical procedure discussed above for the poly-Weibull distribution. The maximum likelihood parameter estimates are shown in Table 11.

Table 11 – Maximum likelihood estimates for the censored Meeker data set

Model	MLE of the Parameters					
Tri-Weibull	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
	6.795	0.742	22.125	338.686	346.727	974.718
Bi-Weibull	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$		
	6.795	0.742	338.686	346.727		
EMWE	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\gamma}$		
	93.100	2.2 x 10^{-3}	0.633	0.625		
NMW	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	
	0.0142	0.0346	0.6939	0.3015	0.0117	

Here, $\hat{\beta}_1, \hat{\beta}_2, \hat{\alpha}_1, \hat{\alpha}_2$ are equivalent for the bi-Weibull and tri-Weibull distributions as $\hat{\beta}_3$ and $\hat{\alpha}_3$ provide no additional information, resulting in a singular Hessian for the tri-Weibull. Thus, the goodness of fit measures displayed in Table 12 show that bi-Weibull is the preferred model.

Table 12 – Performance measures for the censored Meeker data set

Model	Params	Log-Lik	K-S	p-value	AIC
Bi-Weibull	4	-140.95	0.1304	0.9924	289.9
NMW	5	-141.16	0.1304	0.9924	292.5
Tri-Weibull	6	-140.95	0.1304	0.9924	293.9
EMWE	4	-151.43	0.2609	0.4218	310.9

Figure 15 shows that the poly-Weibull and NMW fit the data equally well, while the performance of the EMWE reflects convergence problems associated with the model. Note that in each plot the bi- and tri-Weibull curves are coincident.

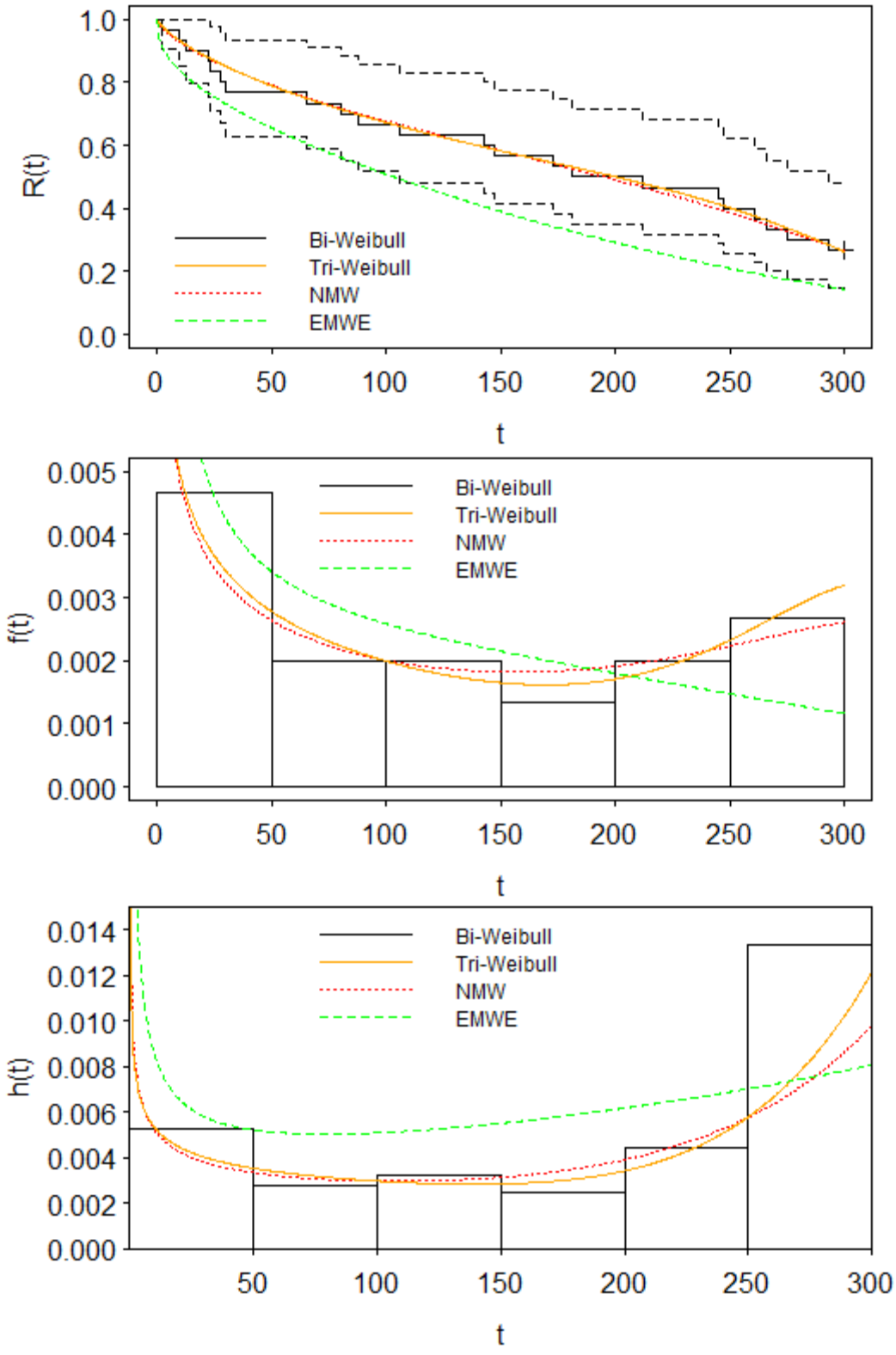

Figure 15 – Triple plot illustrating the fit of the bi-Weibull, tri-Weibull NMW and EMWE models to the censored Meeker data

Conclusion

The poly-Weibull distribution was presented as an attractive alternative to the standard two-parameter Weibull for fitting data sets where the hazard function is monotone, non-monotone or even bathtub shaped. Numerical and analytical procedures for obtaining the poly-Weibull maximum likelihood parameter estimates and asymptotic standard errors are presented. It was shown that the equations for obtaining the poly-Weibull maximum likelihood parameter estimates initially derived and presented by Davison and Louzada-Neto contained an error. The corrected equations were therefore derived and verified along with the components of the observed Fisher Information matrix for the generalized poly-Weibull distribution. The goodness of fit for two forms of the poly-Weibull distribution, the bi-Weibull and tri-Weibull, was then assessed for two reference data sets using the Akaike information criterion and Kolmogorov-Smirnov test statistic. Our results show that the bi-Weibull and the tri-Weibull outperform two other modified Weibull distributions with respect to their fit of data for which the hazard rate function is bathtub shaped. Further, the tri-Weibull dominates the other models when the bathtub shape is prominent. However, this advantage is overcome by the penalty incurred in the AIC from additional parameters when the bathtub shape is less apparent. Thus it is our recommendation that the tri-Weibull distribution be considered when one needs to model data with a pronounced bathtub shaped hazard function. When the data does not reflect a prominent bathtub shape the bi-Weibull distribution is an excellent choice.

V. Characterizing Reliability Growth in Early System Design

Jason K. Freels
Joseph J. Pignatiello
Richard L. Warr
Raymond R. Hill

ABSTRACT

A successful reliability growth program requires setting and achieving reliability targets throughout the various stages of a system's development process. However, schedule and cost constraints often preclude sufficient testing on early prototypes to generate meaningful reliability estimates. While qualitative accelerated test methods quickly improve system reliability by identifying and removing initial design flaws, no attempt is made at quantifying the reliability. In the current paper a modified accelerated life test is proposed whereby the projected system reliability can be estimated after implementing corrective action. Assuming the system contains an unknown number of independent competing failure modes whose respective time to occurrence is governed by a distinct Weibull law, the observed failure times are modeled with the poly-Weibull distribution. We show that under a qualitative accelerated life test scenario the poly-Weibull failure process can be modeled with a Weibull distribution for various sample sizes and system types. Thus, the proposed model utilizes the Weibull distribution to estimate system reliability after one or more failure modes have been discovered and removed. To our knowledge the proposed model is the first attempt to incorporate physical acceleration models into reliability growth planning and is intended to serve as a prototype upon which refinements in reliability growth modeling can be incorporated.

Introduction

Reliability growth testing has typically existed within the purview of large development efforts, such as military weapon systems [77] where failures may not be observed until the system is inspected at the end of a testing phase. In this context, the metric of interest for many reliability growth models, such as those based on the power law process (PLP) [39], is the cumulative mean time between failures within a given phase [9]. Accordingly, these models express reliability growth as a change in the rate of occurrence of failures (ROCOF) across subsequent

test phases. But reliability growth testing can take several months or longer to reach a reliability goal with a specified level of confidence as PLP-based models assume the test conditions reflect the intended usage environment. Consequently, accelerated testing methods have been viewed as a means to reduce the duration of a reliability growth test.

Introducing accelerated stresses in a reliability growth test can increase the ROCOF as the energy put into the system is sufficient to activate failure modes that would otherwise remain dormant under use-level stress. However, no model exists to describe how changes in the level of an accelerating variable affect the ROCOF. Moreover, when multiple failure modes are exposed in a test, separate models are required to characterize the life-stress relationship for each mode [54] as models describing the behavior of entire systems under accelerated stress also do not exist.

One particular testing strategy, qualitative accelerated reliability testing [12, 11, 13], exposes early product designs to elevated environmental and repetitive use stresses to force latent defects to become manifest thereby ensuring that with appropriate mitigation a reliable product is fielded quickly. However, qualitative test methods are designed to discover and remove failure modes quickly, without regard for reliability estimation and many tests employ a test-fix-test strategy where only a single observation of each failure mode is obtained prior to its removal through corrective action. Test planners have long sought a way to incorporate the results of qualitative accelerated testing into an estimate of system reliability in the early design and development phases. While case-studies [15, 16, 14] attest to the

effectiveness of qualitative testing for producing reliable products, translating the limited failure data into a meaningful measure of reliability growth has remained elusive. Unlike qualitative testing which discovers previously unknown failure modes for which little information on the mode-specific time to failure distributions may be available, quantitative accelerated life tests (QALT) investigates the probability of failure due to a few known failure modes.

In a quantitative test, prior knowledge of the relationship between the failure mechanism and the accelerating variable is required to identify an appropriate life-stress model used to describe how the time to failure distribution changes with increased stress [54, 45]. For qualitative tests, prior knowledge is unavailable since the failure modes may be unknown prior to their occurrence and the stresses used to discover the failures often follow complex profiles [17]. Further, corrective action alters the system design configuration by eliminating the flaws causing failure. Since subsequent configurations represent different systems, their respective failure data cannot be combined in a meaningful way. To address this issue the current paper proposes a modified qualitative test and an associated model whereby observations obtained in an accelerated test can be used to estimate the projected system reliability after corrective action.

Testing Framework

In the proposed testing scenario a sample of n prototypes, sharing a common initial design configuration are exposed to a constant elevated stress state until failure (i.e., there is no censoring). The stress state may be comprised of one or

more stressors, but the levels of each stressor must remain constant throughout the test sequence until all n prototypes have failed. Once the test is complete, corrective action is implemented on all of the prototypes in the sample set to remove each distinct failure mode discovered thereby restoring the prototypes to common design configuration. In subsequent tests, the levels of one or more stressors may be changed to create a more extreme stress state capable of generating additional failures. This process continues until a failure occurs that is deemed too cost prohibitive to remove at which point the test is concluded and the design configuration is frozen. In the next section we introduce the model, first discussing the analysis of data from an accelerated test in which each system contains a single failure mode and no corrective action is implemented. Next, the data are re-analyzed assuming that corrective action is implemented after the test. Finally, we consider the general case where multiple failure modes exist in each system and corrective action is implemented.

Model

Single Failure Mode Case – No Corrective Action

In the simplest case, only one failure mode exists in each system. For this case the n times-to-failure can be modeled by the distribution function $F(t;\theta)$ where the parameter vector is a function of stress, $\theta(s)$. In the analysis of accelerated test data it is often assumed that the relative standard deviation ($\sigma \cdot \mu^{-1}$) of the failure observations is unaffected by changes in the applied stress [100]. For shape-scale distributions such as the Weibull(α,β), gamma(α,β), and log-normal(e^{μ},σ^2) this

implies that the functional relationship between a product's lifetime distribution and the applied stress is reflected in the scale parameter only, while the shape parameter is independent of stress.

When the failures from an accelerated life test are due to a single failure mode the observations can be rank ordered from smallest to largest and plotted against their corresponding quantiles as derived from a nonparametric estimate of the cumulative distribution function, $\hat{F}(t)$. This nonparametric estimate can be found using a number of plotting position models [101, pp. 6-8], such as the median plotting position

$$\hat{F}(t_i) = 100 \cdot \left(\frac{i - 0.3}{n + 0.4}\right) \tag{38}$$

where i denotes the i^{th} ordered failure and n represents the total number of observations. To determine if the data follow a specific candidate distribution, the plotting coordinate values and failure times are substituted into the "linearized" distribution function. If the data are derived from the candidate model the transformed observations should closely trace a line. Using the Weibull(α, β) distribution as an example, the linearized distribution function is found by rewriting the CDF as

$$\log_{10}\left[\ln\left(\frac{1}{1 - \hat{F}(t_i)}\right)\right] = \beta \log_{10}(t_i) - \beta \log_{10}(\alpha). \tag{39}$$

If the data are consistent with the Weibull model, the resulting plot of $-\ln\left(1 - \hat{F}(t_i)\right)$ versus t_i will be nearly linear with slope β when plotted on a log-log scale (Figure 16). Alternatively, one may also use specially designed plotting

papers that have been developed for several distributions where the axes have been transformed such that the untransformed failure times and non-parametric probability estimates $\hat{F}(t_i)$ may be plotted directly. If Weibull plotting paper is used α and β may be estimated graphically, observing where the fit line crosses the line designating the 0.632 quantile and the slope of the plot. Figure 16 displays the raw data and linear fits from an accelerated life test discussed in [54, p. 115] where the results are hours to failure obtained at three levels of temperature. Because $\beta_{220} = \beta_{240} = \beta_{260} = \beta$ was assumed, the fit lines should be nearly parallel. Kececioglu [55] notes that for small data sets ($i.e.$, those where n < 10 samples) the slope of the distribution fit lines can vary widely from their true values. Nelson [54] suggests that unless the line slopes change systematically with stress or the slope for one stress level is dramatically different than the others, the lines may be redrawn parallel to fit with the constant β assumption. From Figure 16 it is clear that the slopes of the observations obtained at 220^oC and 240^oC are nearly parallel and the slope of the 260^o data is not dramatically different. Thus the 260^oC fit line is redrawn such that all three fit lines are parallel and the median time to failure is unchanged as shown in Figure 17. From these adjusted plots it is determined that $\beta = 6$ and $\alpha_{220} = 2850, \alpha_{240} = 1700, \alpha_{260} = 1040$.

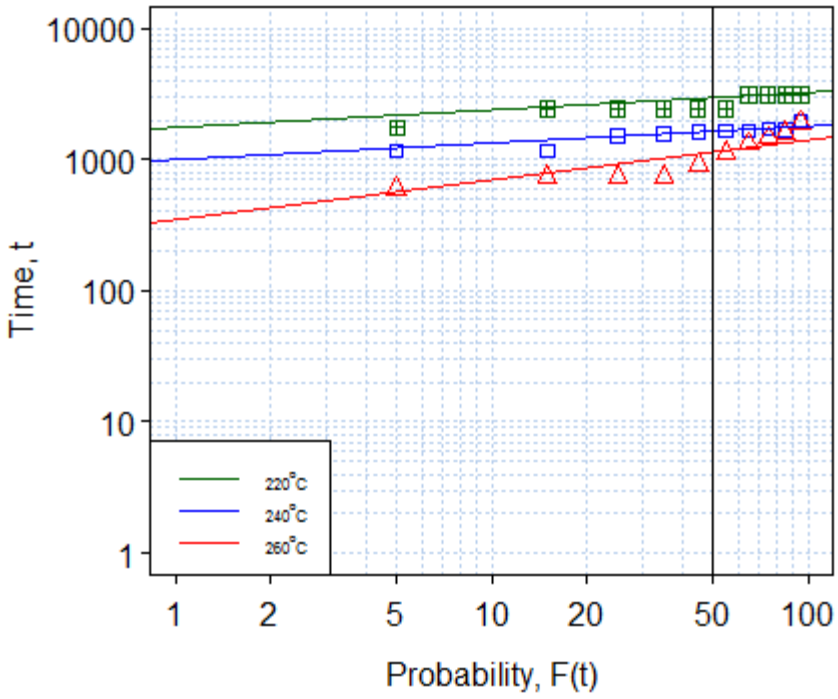

Figure 16 – Weibull plot of failure times observed at three levels of temperature in an example accelerated life test [18]

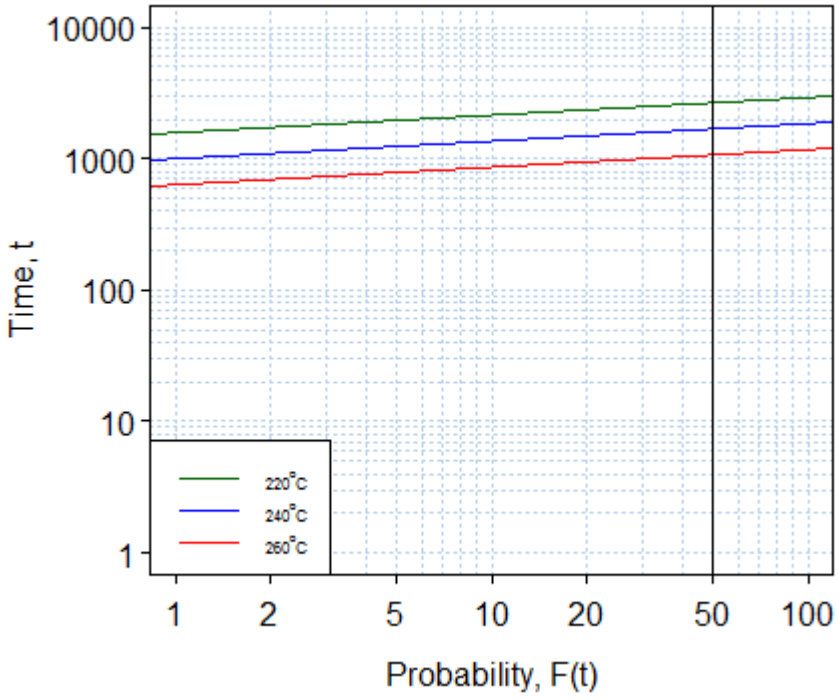

Figure 17 – Weibull plot of example failure data [18] redrawn parallel under the assumption that the shape parameter is independent of stress

The purpose of the test discussed in [54] was to determine if the design requirement of a 20,000 hour median life at the design temperature (180^0C) could be met. To make this determination, the median life at each temperature is found by noting where the black vertical lines in Figures 16 and 17, denoting the median failure probability, crosses each fit line. These median values are used to construct a relationship plot (Figure 18) showing how the median life changes with increasing temperature. Because temperature is the accelerating variable, the Arrhenius reaction rate [45] equation was used to model the life-temperature relationship. The Arrhenius life-stress model is given by

$$L(T) = C\exp(AT^{-1}) \tag{40}$$

where $L(T)$ represents a quantifiable life measure such as characteristic life or median life, A and C are model parameters to be determined, and T is the absolute temperature in degrees Kelvin. The parameter C is related to the specimen geometry and testing methodology, while A is a function of the activation energy required for a reaction to proceed [54]. To determine if the Arrhenius model accurately describes the life-temperature relationship exhibited by the data, a process similar to what was described for fitting candidate distributions to the raw data at each temperature can be employed. The Arrhenius model is first "linearized" to obtain

$$\ln[L(T)] = A(T^{-1}) + \ln(C). \tag{41}$$

If the Arrhenius model is correct, the $\ln[L(T)]$ values plotted against the reciprocal of the absolute temperature on standard plotting paper will approximately fall on a

line with slope A and intercept $\ln(C)$ for each set of temperature data. Upon analyzing the data, the Arrhenius parameter values for the plot in Figure 18 were estimated to be $\hat{A} = 6640.1$ and $\hat{C} = 4.055 \cdot 10^{-3}$.

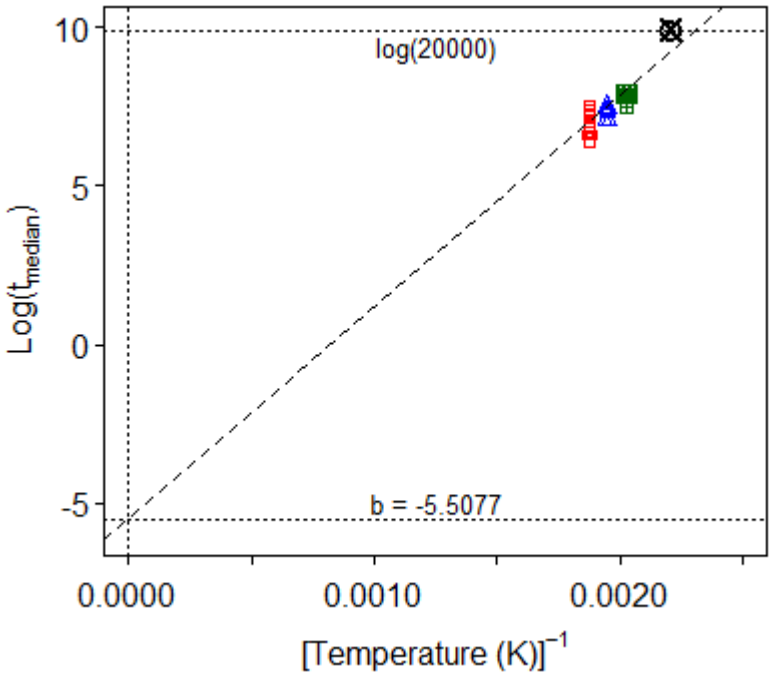

Figure 18 – Arrhenius relationship plot of the example data

Alternatively, the data may be presented with respect to temperature, rather that reciprocal temperature as shown in Figure 19. In this figure the median times to failure (solid line) are plotted along with the 0.1 and 0.9 quantile times to failure against the temperature in degrees Celsius. Extrapolating the solid line in this figure back to the use temperature of $180^{o}C$ clearly shows that the current median life of 9000 hours falls well short of the 20,000 hour design requirement.

In the next section we revisit this example, utilizing the proposed approach in which corrective action is implemented after testing is performed at each

temperature level. The impact of implementing corrective action is investigated assuming the times to failure observed at $220^{o}C$ in this second iteration are identical to the results discussed above. A benefit of the proposed approach is that the number of required samples can be reduced since corrective action will allow the samples to be reused at each temperature level.

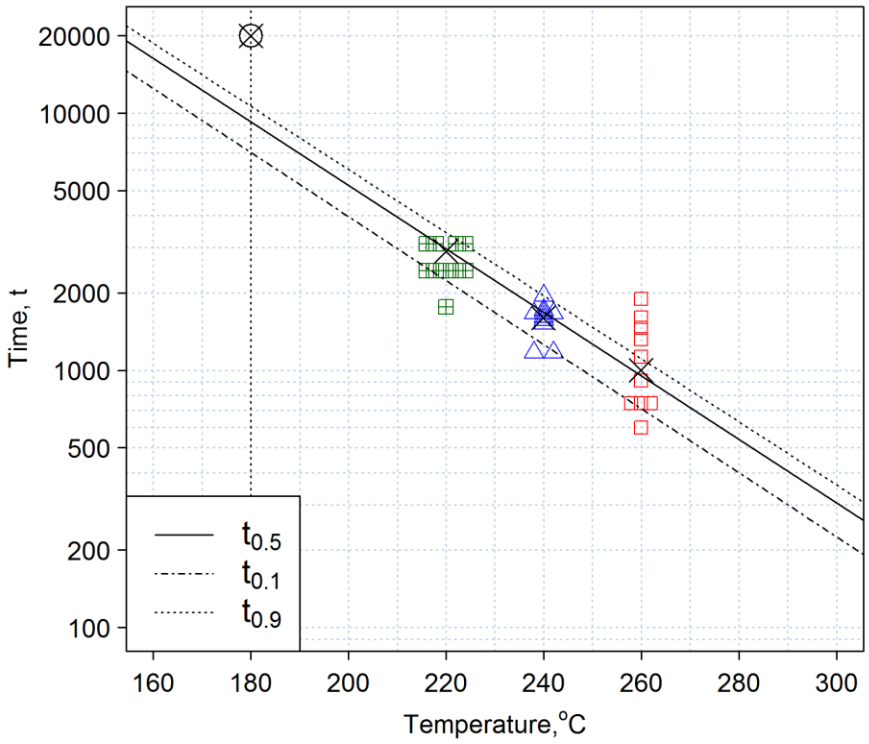

Figure 19 – Adjusted Arrhenius relationship plot of example data versus temperature

Single Failure Mode Case – With Corrective Action

After a test, each failure is investigated to determine its root cause and an appropriate corrective action strategy is then designed to eliminate the failure in subsequent tests. However, corrective action rarely eliminates a failure mode

completely. It is common in the reliability growth literature [9] to characterize the impact of a corrective action by modeling the reduction in a failure mode's hazard rate by a proportional amount $\rho \in (0,1)$ known as the fix effectiveness factor (FEF). The FEF varies for each failure mode, but [9, p. 29] reports that the typical range of values for many government and industry systems is $(0.55, 0.85)$ with a typical value of 0.70. The FEF value for a failure mode is often subjectively estimated after an extensive scoring process in which design engineers and failure analysis experts consider various corrective action alternatives. As a result of implementing corrective action the system reliability improvement can be expressed as

$$\frac{h'(t)}{h(t)} = 1 - \rho \tag{42}$$

where $h'(t)$ and $h(t)$ denote the hazard rate functions of the corrected and uncorrected systems, respectively. For failure modes with Weibull distributed times to occurrence, the reduction in hazard rate resulting from corrective action with fix effectiveness factor ρ can be represented as

$$1 - \rho = \frac{h'(t)}{h(t)} = \frac{\beta t^{\beta-1}}{(\alpha')^{\beta}} \left(\frac{\beta t^{\beta-1}}{(\alpha)^{\beta}} \right)^{-1} = \left(\frac{\alpha}{\alpha'} \right)^{\beta}. \tag{43}$$

Substituting the Arrhenius model for the Weibull scale parameter in (47) results in

$$1 - \rho = (Ce^{A}/Ce^{A'})^{\frac{\beta}{T}} = [\exp(A - A')]^{\frac{\beta}{T}}. \tag{44}$$

The simplification in (44) is made under the assumption that the parameter C is unchanged with corrective action. Thus, the reduction in failure rate resulting from a corrective action with fix effectiveness ρ can be attributed to an increase in the

activation energy required to initiate the failure process, where the increased activation energy is expressed as

$$A' = A - \left(\frac{T}{\beta}\right)\ln[1-\rho]. \tag{45}$$

However, (45) implies that the value of A' resulting from corrective action depends on the temperature at which the failure was discovered. While increased temperatures allow failures to be observed more quickly, the strategy to remove the failures is independent of temperature, thus there exists a unique value of A' for each corrective action. In traditional reliability growth models, the estimated value of ρ is based on an assumed mission profile of the fielded system [9, p. 2]. Introducing accelerated stresses to a reliability growth test does not effect this interpretation of ρ and the improved activation energy is expressed as

$$A' = A - \left(\frac{T_{use}}{\beta}\right)\ln[1-\rho]. \tag{46}$$

Table 13 displays the projected reliability measures at the use stress of $180^{o}C$ that result from implementing three levels of corrective action after the first test sequence.

Table 13 – Projected reliability measures at $180^{o}C$ after corrective action

$FEF\ (\rho)$	α'	t'_{med}	$A' - A$
0.85	12855.95	12094.14	143.28
0.70	11453.35	10774.65	90.93
0.55	10704.93	10070.59	60.31

Figure 20 displays the Arrhenius plots corresponding to the projected reliability measures in Table 13 along with the original plot of the uncorrected system. Each

plot in the figure has common intercept $\ln[C]$ and slope increased by $A' - A$ over the uncorrected system. The vertical line in the figure denotes the use temperature $180^{o}C$. The ordinate of this line is marked to indicate the 20,000 hour requirement. It can be seen in Figure 20 that none of the projected Arrhenius plots corresponding to the three FEF values will enable the system to meet the 20,000 hour requirement.

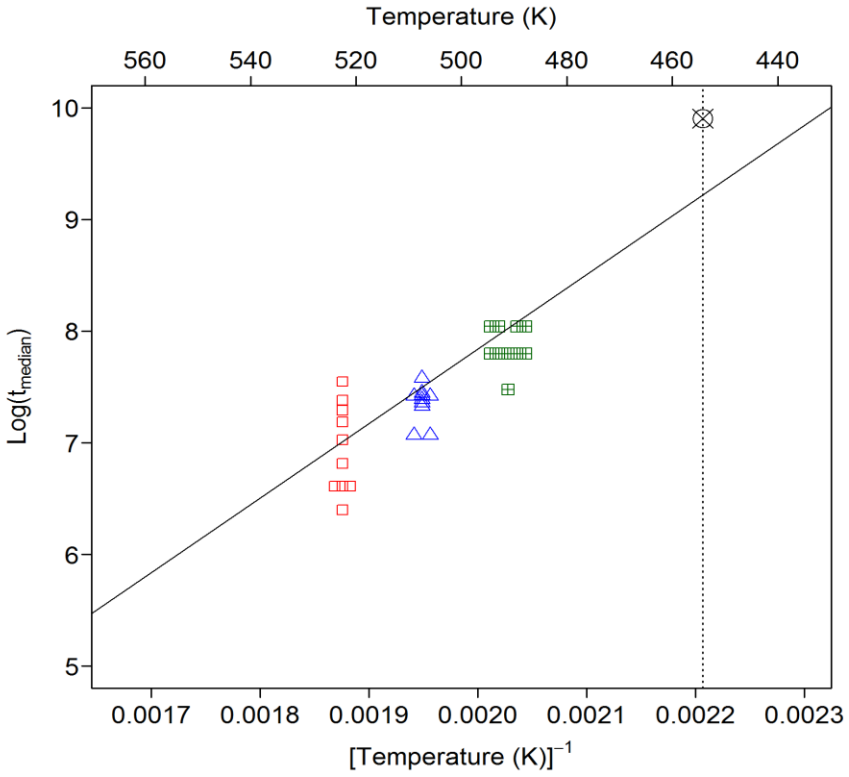

Figure 20 – Updated Arrhenius plot of example data after corrective action

This same conclusion could also be drawn by observing the projected failure distributions associated with each of the three levels of fix effectiveness ρ as shown in Figure 21. The projected distribution function $WEIB(\alpha'\beta)$ resulting from corrective action is expressed as

$$P(X \leq x;T) = 1 - \exp\left(\frac{x}{C\exp\left[\left(\ln\left[(1-\rho)^{-\frac{T_{use}}{\beta}}\right] + A\right)T^{-1}\right]}\right)^{\beta}. \qquad (47)$$

where T represents any temperature of interest. Figure 21 presents the density and reliability functions for the uncorrected system at $180^{o}C$ along with the projected density and reliability functions after corrective action. The vertical lines spanning both plots in the figure represent the median values for each distribution. Again, it is clear that a single corrective action is insufficient to achieve the requirement as the maximum projected median life is just over 12,000 hours.

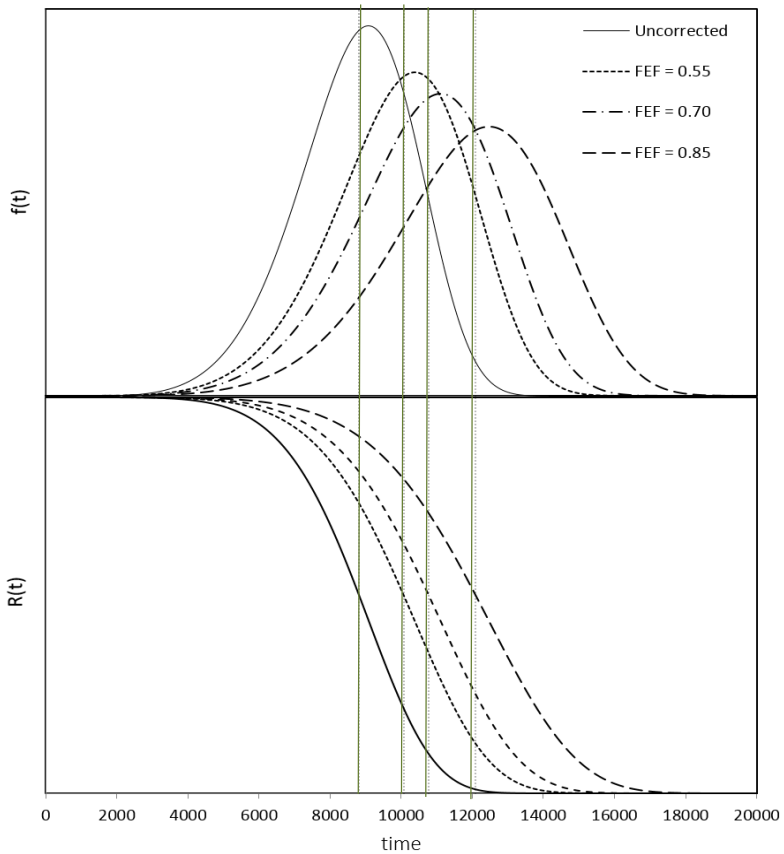

Figure 21 – Joint plot of density and reliability functions at $180^{o}C$ with corrective action

Assuming corrective action had been implemented to reduce the intensity of the failure mode discovered during the $220^{o}C$ test, the improved system would be less sensitive to elevated temperatures. Quickly discovering additional failures in the improved system would therefore require the application of a more severe environment than was used to identify the failure mode initially. For systems with a single failure mode, the fix effectiveness of a second corrective action may be far lower than that of the initial corrective action. Assuming k corrective actions are feasible system reliability can be projected recursively and the Arrhenius line slope after the k^{th} corrective action becomes

$$A^{(k)} = \sum_{k=1}^{K} \left(\ln \left[\left(1 - \rho^{(k)}\right)^{-\frac{T}{\beta}} \right] \right) + A \tag{48}$$

and the projected distribution function of system life at T_{use} is expressed as

$$P(X \leq x; T) = 1 - \exp \left(\frac{x}{C \exp \left[\left(\sum_{k}^{K} \left(\ln \left[(1 - \rho_k)^{-\frac{T_{use}}{\beta}} \right] \right) + A \right) T_{use}^{-1} \right]} \right)^{\beta}. \tag{49}$$

A final analysis of the test scenario discussed in [54] revealed that attaining the 20,000 hour median failure time at the use stress of $180^{0}C$ would require a 97% reduction in the initial hazard rate. A reduction of this magnitude would correspond to implementing three corrective actions with each corrective action having a very high fix effectiveness factor of 0.85. This unlikely event confirms the additional commentary provided by Nelson [54, p. 403] where further testing on a subsequent design indicated that it was still insufficient to attain the requirement and the project was abandoned. Using the proposed model would have provided an

indication of the risk in attaining the required reliability while utilizing less than half of the specimens.

Design of Accelerated Reliability Growth Tests

The typical objectives in designing an accelerated reliability growth test are to minimize the total test time, the number of required samples and the prediction variance of the estimate of system life in the usage environment. Because qualitative testing is conducted during the early phases of product development, the number of prototypes available for testing is already severely limited due to the cost of production. Thus, identifying a best qualitative test design is reduced to selecting the stress levels at which the accelerated test will be performed and how the available samples will be allocated to those stresses.

To model lifetime data from an accelerated test, it is generally assumed that the log of the failure times observed at each stress level follow a location-scale distribution

$$P(Y \leq y) = \Phi\left(\frac{y-\mu}{\sigma}\right) \tag{50}$$

where μ and σ are the true but unknown location and scale parameters and $\Phi(\cdot)$ denotes the standard form of the location-scale distribution. Further, it is assumed that the value of the scale parameter σ is independent of stress while the mean log-failure time is represented as a linear function of the coded stress level x_j expressed as

$$\mu(x_j) = c_o + c_1 x_j. \tag{51}$$

For the Arrhenius relationship, the coded stress level is $x_j = B/T_j$ where B is Boltzmann's constant and T_j refers to the temperature in an absolute scale such as Kelvin or Rankine. The variance of the estimate of system life at the use-level stress level x_0 as result of observing failures at stress level $x_i, i = 1, 2, \dots, I$ is expressed as

$$Var[m(x_0)] = \left\{1 + (x_0 - \bar{x})^2 \left[\frac{n}{\sum(x_i - \bar{x})^2}\right]\right\}\left(\frac{\sigma^2}{n}\right) \tag{52}$$

where $m(\,)$ denotes the life measure of interest, such as the mean or median, and n is the number of specimens in the sample. The variance in (52) is minimized by allocating the n specimens to two stress levels, a high stress level x_H and a low stress level x_L where the high stress level is defined as the upper limit stress level that a specimen can withstand such that the failure modes are representative of the type of failures expected to occur in the usage environment. The use stress level x_0 and high stress level x_H establish the bounds for the design region while the low stress level x_L is chosen to minimize the variance in (52). To determine the optimal value for x_L in a general design region Nelson [54] defines the extrapolation factor ξ_i

$$\xi_i \equiv \frac{x_H - x_i}{x_H - x_L}. \tag{53}$$

At the use stress level x_0, ξ_0 represents the ratio of the allowable design range $x_H - x_0$ to $x_H - x_0$ the design region selected for the accelerated test. It can be shown that the variance expression in (52) can be simplified as

$$Var[m(x_0)] = \left[1 + \frac{(\xi_0 - p)^2}{p(1-p)}\right]\left(\frac{\sigma^2}{n}\right) \tag{54}$$

when two test stresses are chosen that lie at the extremes of the design region, that is x_H and x_L, where p denotes the proportion of specimens allocated to the low

stress level x_L. Further, the optimal allocation of specimens which minimizes the variance in (54) for a general design region is denoted by p^* and expressed as

$$p^* = \frac{\xi_0}{2\xi_0 - 1}. \tag{55}$$

Substituting p^* into (54) results in the minimum prediction variance of the estimate of system life solely as a function of the extrapolation factor

$$Var^*[m(x_0)] = 1 + 4\xi_0(\xi_0 - 1)\left(\frac{\sigma^2}{n}\right). \tag{56}$$

In the extreme case of $x_L = x_0$, both ξ_0 and p^* equal one regardless of the value of x_H indicating that the variance is minimized when all n tests are conducted at x_0. However, such a design would also result in the maximum possible test duration. As x_L increases, the optimal allocation of samples at the use stress, p^*, decreases and reaches a minimum value of 0.5 as $\xi_0 \to \infty$. Clearly, the total test time is minimized by raising x_L and x_H to their highest possible levels. As x_H and x_0 converge, however, achieving a significant reduction in test time requires pushing x_H closer to the destruct limit. Thus care must be taken to ensure that the failures observed at the high stress are representative of the type of failures that may be observed in the use environment. For this reason the two-stress test design may not be robust to the modeling assumptions and the addition of a third stress level may be beneficial. Figure 22 plots the log variance as a function of the extrapolation factor ξ_0 for the optimal two-stress test design along with the two- and three-stress level designs with an equal allocation of samples at each stress level.

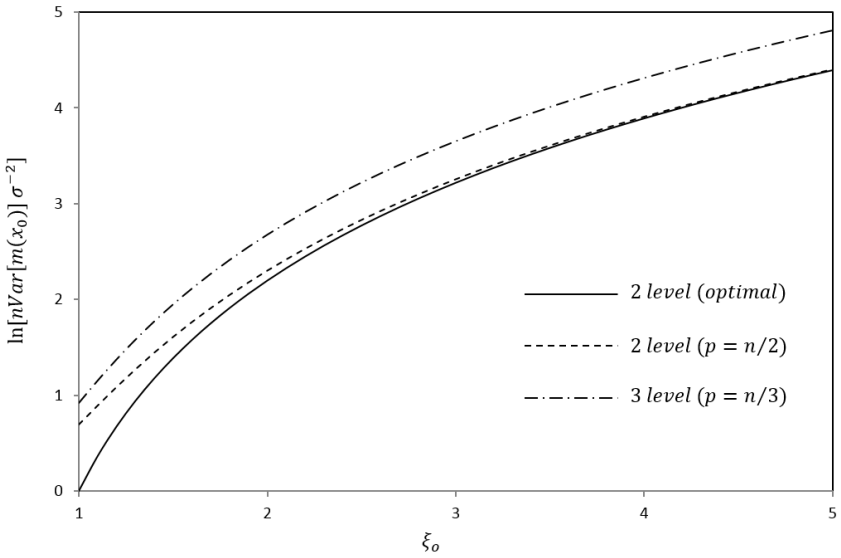

Figure 22 – Variance comparison of the two– and three–stress accelerated test designs with equal allocation against the optimal two-stress test design

When the allowable design region $x_H - x_0$ is large, greater freedom exists to select x_L such that $Var[m(x_0)]$ can be minimized while simultaneously reducing the total test time. Figure 23 shows plots of the total test time against the x_H/x_0 ratio for four values of ξ_0. For each plot, $x_0 = 180^o C$, as was the case in the example test scenario previously described. In each plot the low stress level x_L increases or decreases commensurate with x_H, such that an increase in x_H reduces the time to failure for all n units in the sample.

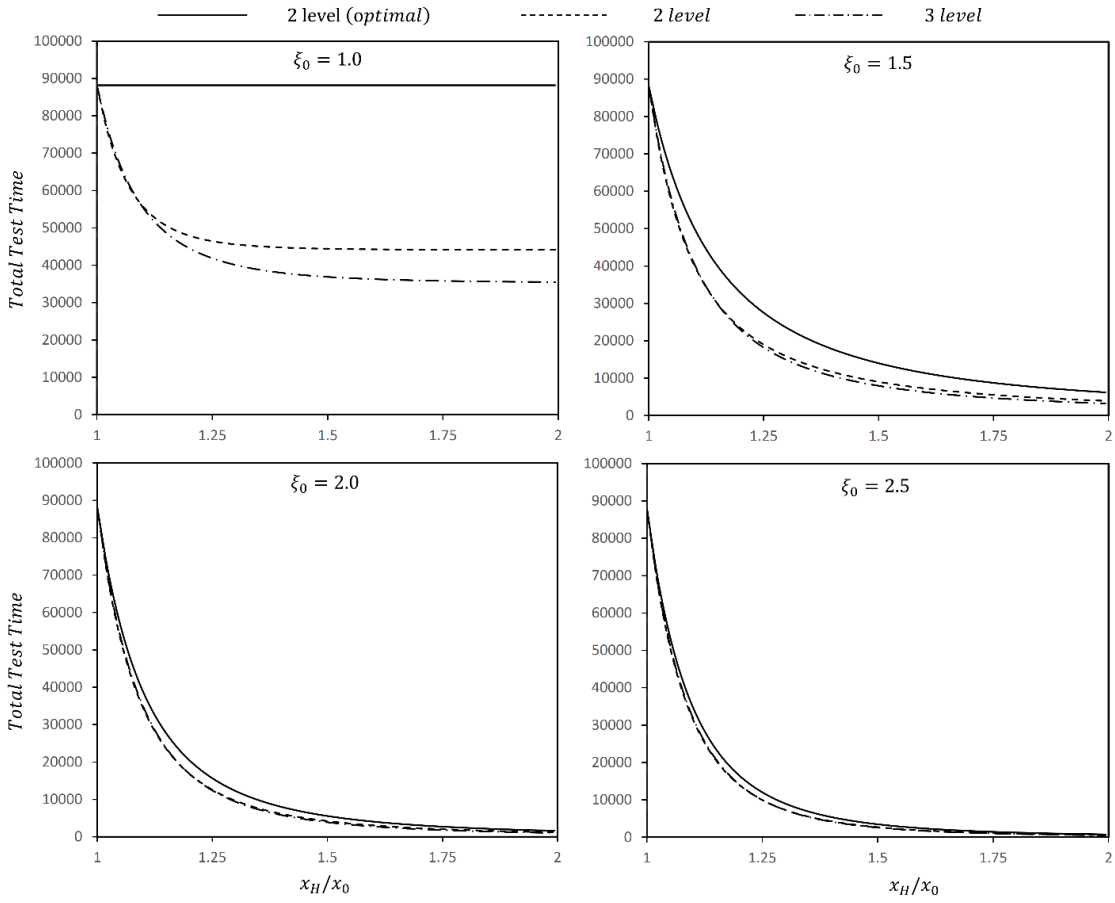

Figure 23 – Comparison of the reduction in total test time for the two-stress equal apportionment design the two-stress optimal design and the three-stress equal apportionment design for various levels of extrapolation factor

Identifying a best design for a given test scenario with inputs x_0 and x_H could be determined through multicriteria optimization where the objective is to simultaneously minimize $Var[m(x_0)]$ and the total test time subject to the constraint that the proportions of samples tested at each stress level sum to one. Figures 22 and 23, indicate that the three stress level design is optimal with respect to minimizing the total test time but sub-optimal for minimizing the variance compared to the two-stress level designs for all values of ξ_0. Alternatively, an optimal design minimizes the asymptotic variance, but maximizes the total time on

test. Therefore, the two-stress equal apportionment design is recommended as a tradeoff between the two-stress optimal design and the three-stress equal apportionment design.

Multiple Failure Modes Case – With Corrective Action

For systems containing multiple independent failure modes, estimating system reliability from accelerated test data requires sufficient samples to observe each failure mode at multiple stress levels [54]. Separate models, similar to those discussed for the single failure mode case, may then be developed for each distinct failure mode and the exact reliability for the overall system can then be determined as

$$R(t) = \prod_{j=1}^{J} R_j(t), \quad j = 1, 2, \ldots, J. \tag{57}$$

In qualitative testing, however, it is rare that a sufficient number of prototypes can be made available to account for all of the failure modes that have yet to be discovered in early system testing. For tests conducted with limited sample sizes, an estimate of system reliability may be obtained by combining the observations of multiple failure modes to form a single distribution plot as was shown above for systems with only one failure mode. In this scenario a possibly unknown number of independent flaws, denoted by J, compete to be the cause of system failure (Figure 24). The observed lifetime for prototype i is therefore represented as the minimum occurrence time among the J modes in the system.

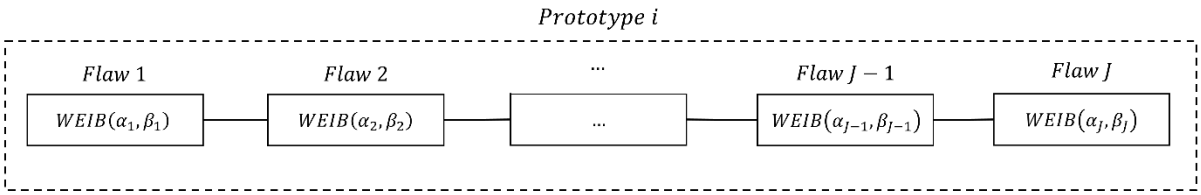

Figure 24 – Serial arrangement of flaws within the prototypes subjected to qualitative accelerated reliability test. The arrangement demonstrates the competing risk assumption in the model where the time to failure for prototype *i* is the minimum activation time among the *J* flaws.

With respect to flaw $j \in \{1, \dots, J\}$ the probability of survival up to time t is represented by the random variable T_j with sub-survivor function $S(j,t) = \Pr(Cause = j, Time > t)$. The sub-survivor function is not a proper reliability function in that $S(j,t) \neq P(T > t|C = j)$. Instead the true reliability function specific to mode j is expressed as

$$R(j,t) = \frac{S(j,t)}{q_j} \tag{58}$$

where $q_j = P(C = j) = F(j,\infty) = R(j,0)$ subject to the constraint $\sum_{j=1}^{J} q_j = 1$.

The form of $R(j,t)$ in (58) may be unknown prior to testing and is often assumed based on the testing scenario or prior knowledge. McLean [80] assumed exponentially distributed occurrence times for each failure mode on the basis of mathematical simplicity. A more general assumption, suggested for the proposed model, is $R(j,t) \sim WEIB\left(\alpha_j, \beta_j\right)$. For such cases where $J \geq 2$, the observed failure time $X_i = \min\left[T_j\right]$ follows a poly-Weibull [85] distribution with vector-valued parameters $\boldsymbol{\alpha}$ and, $\boldsymbol{\beta}$ and density function expressed as

$$f_{PW}(t|\boldsymbol{\alpha},\boldsymbol{\beta}) = h(t)R(t) = \left\{\exp\left[-\sum_{j=1}^{J}\left(\frac{t}{\alpha_j}\right)^{\beta_j}\right]\right\}\sum_{j=1}^{J}\frac{\beta_j t^{\beta_j-1}}{\alpha_j^{\beta_j}}. \quad (59)$$

The poly-Weibull distribution arises naturally in scenarios of competing risks as it describes the minimum of several independent random variables when each follows a distinct Weibull law. When only two flaws exist ($J = 2$), the distribution is known as the bi-Weibull, and $J = 3$ model is naturally called the tri-Weibull distribution.

An advantage of the poly-Weibull distribution lies in its capacity for modeling not only increasing, constant and decreasing hazard functions but also non-monotone hazard functions [95] such as the bathtub curve (Figure 25). This property is important as non-monotone hazard functions are common in practice where a system may undergo an initial "burn-in" prior to periods of useful life and eventual wearout.

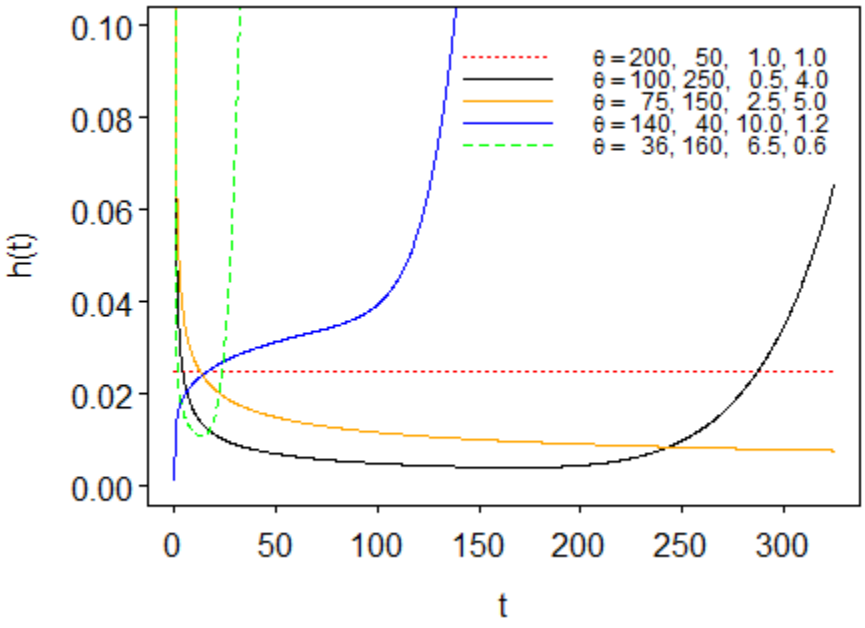

Figure 25 – Example poly–Weibull hazard functions with various parameter values

If the shape parameters $\beta_1, \ldots, \beta_J$ are not sufficiently distinct the poly-Weibull may over-fit the data [27, 28]. For the extreme case where $\beta_1, \ldots, \beta_J = \beta$, the poly-Weibull converges to a Weibull distribution with density $f(t; A, \beta) = \mathrm{A}\beta t^{\beta-1} e^{-\mathrm{A}t^{\beta}}$ and scale parameter

$$\mathrm{A} = \left[\sum_{j}^{J} \alpha_{\mathrm{j}}^{\beta}\right] \cdot \left[\prod_{j}^{J} \alpha_{\mathrm{j}}^{\beta}\right]^{-1}. \tag{60}$$

Our interest is not in estimating α_j, β_j for all $j \in \{1, \ldots, J\}$, as the data produced in qualitative testing is insufficient for this. Moreover, the poly-Weibull CDF cannot be linearized, as was done for the Weibull distribution previously, due to the summation inside the logarithm operator. Thus, the probability plotting procedures discussed above for the single failure mode case cannot be employed for the bi and tri-Weibull distributions. Instead, we demonstrate through simulation that for scenarios specific to qualitative accelerated life testing the performance of Weibull distribution is sufficient for fitting data sets in which the observations represent the minimum time among several competing failure modes each having Weibull distributed occurrence times.

Monte Carlo Simulation Procedure

Prior to a qualitative reliability test, uncertainty exists regarding the number of latent failure modes embedded within each prototype and the expected time to occurrence for each mode. To investigate the validity of the Weibull distribution in the multi-failure mode case, a Monte Carlo simulation study was conducted. The simulation utilized a $5 \times 4 \times 3$ full factorial design with three factors: system

complexity, sample size, and initial system quality as depicted in Figure 26. System complexity is represented by the total number of latent failure modes competing to be the cause of failure in each identical system under test. For this study, five level levels of system complexity ($J = 5, 10, 20, 40, 50$) were included. Next, four levels of sample size ($n = 10, 20, 40, 100$) were utilized to illustrate how the performance of the Weibull distribution compares to the bi- and tri-Weibull as the number of failure observations increases. Finally, the initial system quality is a categorical variable with three levels (low, middle, high).

We suggest that products with high, middle and low levels of initial quality can have an identical number of embedded flaws caused by poor design or manufacturing defects. However, in products with low initial quality a greater degree of variability exists in the severity of each flaw and the mean time until a flaw activates may be far lower in low quality products than for a similar flaw in products with higher levels of quality.

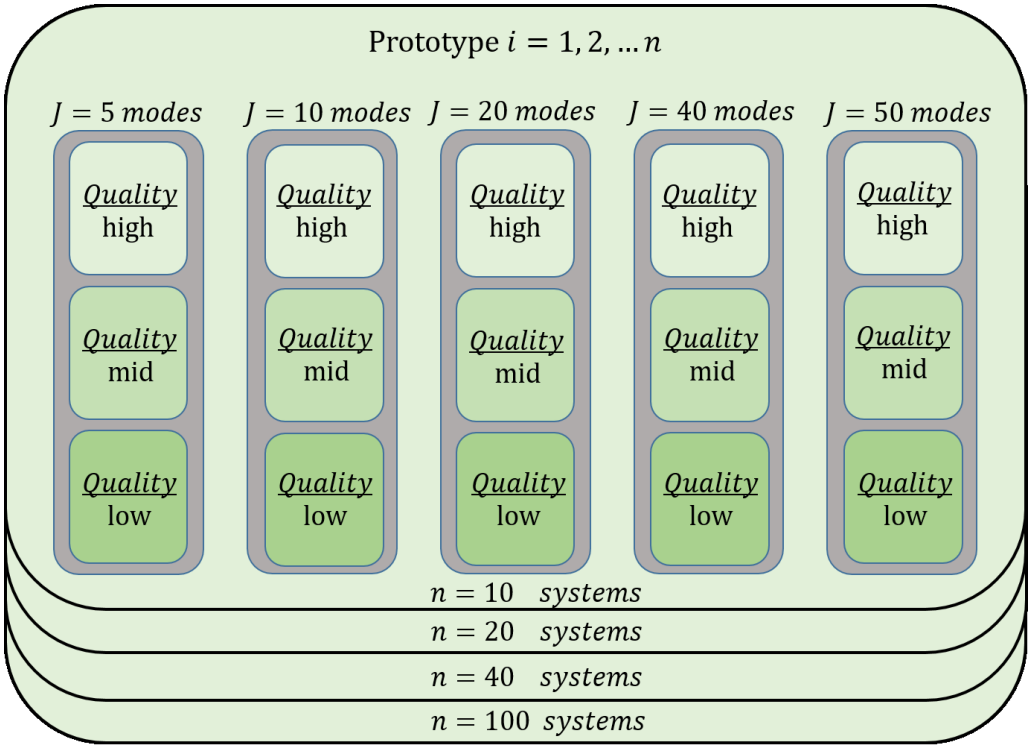

Figure 26 – Graphical depiction of process to simulate observed minimum failure times for systems with varied numbers of competing failure modes and levels of initial design quality

To simulate the time to occurrence distribution for failure mode $j = 1, \ldots, J$ embedded within each of the n prototypes in the sample set, a $J \times 2$ matrix of mode-specific Weibull parameter values was created for each sample. The Monte Carlo approach used for creating these failure mode-specific parameters arises from noting that for random variable $T_j \sim Weib(\alpha_j, \beta_j)$, the coefficient of variation (CoV)

$$CoV[T_j] = \left[\frac{Var[T_j]}{E[T_j]^2}\right]^{1/2} = \left[\frac{\Gamma(1 + 2/\beta_j)}{\Gamma(1 + 1/\beta_j)^2} - 1\right]^{1/2} \quad (61)$$

is solely a function of β_j and is proportional to $1/\beta_j$ as shown in Figure 27. Jin et al. [102] notes that for many electronic systems the CoV tends to vary between 0.05

and 5.50. Thus, an approach for obtaining the mode-specific Weibull shape parameter is to sample a $CoV[T_j]$ value from a distribution, substitute this value into (61), and solve for β_j.

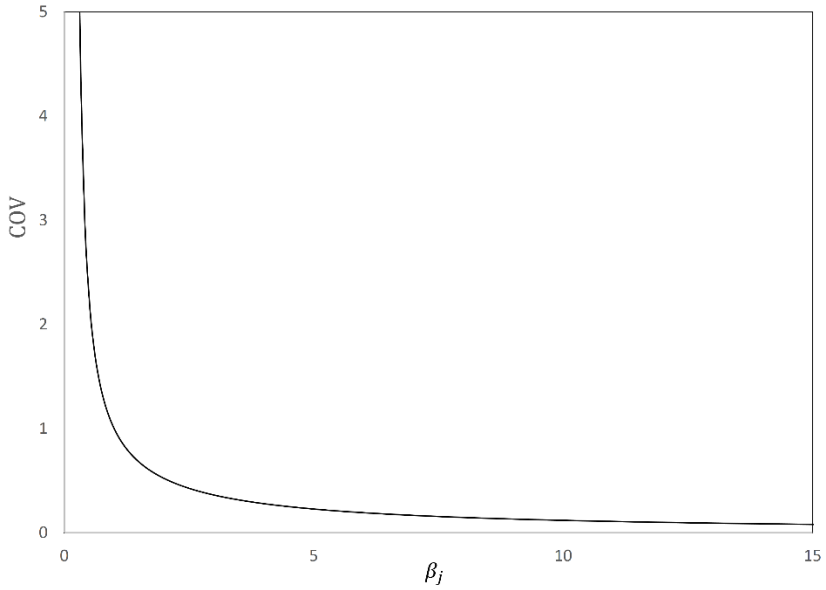

Figure 27 – Relationship between the coefficient of variation and shape parameter B_j for Weibull distributed random variables

The support region for this sampling distribution should mirror the range defined by Jin et al. [102], so for this simulation study a four-parameter beta distribution [103] was used. If $X \sim BETA(\lambda, \theta)$, $x \in [0,1]$ with shape parameters λ and θ, the four-parameter beta results from the transformation $Y = X(d - c) + c$ where d and c are the upper and lower limits of the desired support for the transformed variable Y. Thus, $Y \sim BETA(\lambda, \theta, c, d)$, $y \in [c, d]$ with density function expressed as

$$f(Y; \lambda, \theta, c, d) = \frac{1}{\mathrm{B}(\lambda, \theta)} \cdot \frac{(Y - c)^{\lambda - 1}(d - Y)^{\theta - 1}}{(d - c)^{\lambda + \theta - 1}}. \quad (62)$$

The combination of values for the shape parameters λ and θ can then be used to reflect the levels of design quality that a system may possess at the start of a test. As illustrated in Figure 28, the combination of high variability $Var[T_j]$ and low mean occurrence times $E[T_j]$ for failure modes in low quality systems, imply that the likelihood of $CoV[T_j]$ values at the upper end of the range defined by [102] is greater than in systems with higher quality levels.

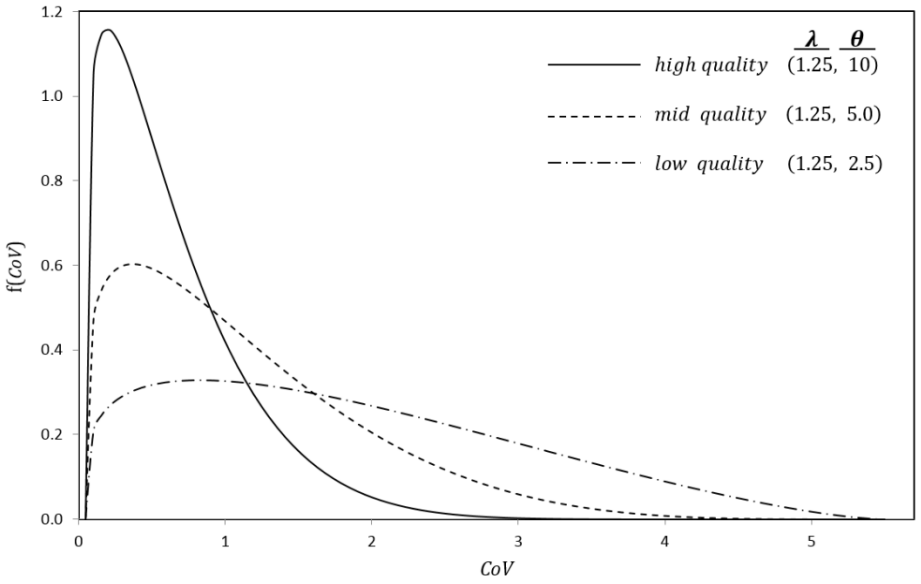

Figure 28 – Plots illustrating the distribution of the coefficient of variation values for multiple levels of initial design quality

In this simulation study the sampling distribution for $CoV[T_j]$ in low quality products had shape parameters $(\lambda, \theta) = (1.25, 2.5)$ to reflect an increased likelihood of an embedded failure mode causing infant mortality. To model products with middle and high levels of initial quality, $\theta = 5$ and 10, respectively, were used. These parameter values result in a sampling distribution for $CoV[T_j]$ that is more

positively skewed than that of low quality systems indicating a reduced likelihood of infant mortality failures.

The Weibull shape parameter obtained from the sampling procedure outlined in the preceding paragraphs may then be used to find the Weibull scale parameter for failure mode j by substituting β_j into the expression

$$\alpha_j = E[T_j] \times \Gamma(1 + 1/\beta_j)^{-1}. \tag{63}$$

As previously noted, higher coefficients of variation imply lower mean occurrence times, therefore the sampling distribution for $E[T_j]$ used in this study is conditioned on $CoV[T_j]$. For this simulation $E[T_j] \sim BETA(\gamma, \delta, 3000, 0)$ where the support region $(0, 3000)$ represents the average annual usage in hours of an electronic device. The corresponding shape parameters γ and δ are determined as

$$\begin{aligned} \gamma &= \left(1 + \frac{CoV_{max} - CoV[T_j]}{CoV_{max} - CoV_{min}}\right)^2 \in [1,\ 4] \\ \delta &= \left(1 + \frac{CoV[T_j] - CoV_{min}}{CoV_{max} - CoV_{min}}\right)^2 \in [1,\ 4]. \end{aligned} \tag{64}$$

Each run in the proposed Monte Carlo simulation can be parameterized by the vector $[n, J, \lambda, \theta, \gamma, \delta]$ designating the sample size, system complexity, and quality level associated with a given test scenario. For each of the 60 scenarios the simulation was iterated 10,000 times. During each iteration, the parameters α_j and β_j, $j = 1, \ldots, J$ were generated and the J Weibull distributions sampled to provide possible failure times. From these J values the minimum is chosen as the

observed failure time for prototype i. This procedure was then repeated for each of the remaining prototypes in the sample.

Using the n failure observations generated in each sample, the relative performance of the Weibull distribution was compared to the bi and tri-Weibull using the corrected Akaike Information Criterion ($AICc$) [104] expressed as

$$AICc = 2k - 2(\mathcal{L}) + \frac{2k(k+1)}{n-k-1} \tag{65}$$

where k denotes the number of parameters in each candidate model and $\mathcal{L}$ is the value of the log-likelihood function evaluated at the maximum likelihood parameter estimates. The MLE's for the bi and tri-Weibull distributions, $\widehat{\boldsymbol{\alpha}}_{MLE}$ and $\widehat{\boldsymbol{\beta}}_{MLE}$, were obtained by solving the system of $2V$ nonlinear equations

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha_v} &= \sum_{i=1}^{n} \left(\frac{t_i}{\beta_v} - \frac{\delta_i}{h(t_i)}\right)\left(\frac{\beta_v}{\alpha_v}\right) h_v(t_i) = 0 \qquad v = 1, \dots, V \\ \frac{\partial \mathcal{L}}{\partial \beta_v} &= \sum_{i=1}^{n} \left[\delta_i \frac{h_v(t_i)}{h(t_i)}\left\{\beta_v^{-1} + \log\left(\frac{t_i}{\alpha_v}\right)\right\} - \frac{t_i}{\beta_v} h_v(t_i) \log\left(\frac{t_i}{\alpha_v}\right)\right] = 0 \qquad v = 1, \dots, V \end{aligned} \tag{66}$$

where $h_v(t_i) = \beta_v t^{\beta_v - 1} \alpha_v^{-\beta_v}$ and $h(t_i) = \sum_{v=1}^{V} \left(\beta_v t^{\beta_v - 1} \alpha_v^{-\beta_v}\right)$. Solving this system of equations cannot be accomplished analytically and numerical root-finding techniques can be tedious as finding a solution can be sensitive on the starting values for the parameters in each equation. We found it simpler to obtain accurate parameter estimates by maximizing the log-likelihood function directly using a quasi-Newtonian method [98] and utilizing (66) to verify the solution. For each simulated data set, the $AICc$ of one of the three candidate distributions has the

lowest value implying that the model corresponding to $AICc_{min}$ has the highest relative likelihood of minimizing the information loss expressed by

$$L(m_i|data) = \exp[0.5(AICc_{min} - AICc_i)]. \quad (67)$$

Appendix A displays histograms of these likelihood values for the Weibull, bi-Weibull and tri-Weibull distributions for each of the sixty testing scenarios.

Next, Anderson-Darling [105] goodness of fit tests were performed to determine the likelihood that each simulated data set could have been drawn from the $WEIB\left(\widehat{\boldsymbol{\alpha}}_{MLE}, \widehat{\boldsymbol{\beta}}_{MLE}\right)$ distribution. The unmodified Anderson-Darling test is performed using the test statistic

$$A^2 = \sum_{i}^{n} \frac{1-2i}{n} [\ln(F(Y_i) + \ln(1 - F(Y_{n+1-i})] - n. \quad (68)$$

where $F(\,)$ denotes the CDF of the distribution of interest. Stephens [106] found A^2 to be one of the best empirical distribution function based test statistics for detecting departures from several continuous distributions, including the Weibull. Because the scale and shape parameters are estimated from the data generated in each iteration of the Monte Carlo simulation, the modified Anderson-Darling statistic [107] $A^* = MA^2$ was used. Stephens [107, p. 146] identified the functional form of M for the cases where one or both of the Weibull parameters are estimated from the data. Substituting the Weibull distribution function and $M = 1 + \sqrt{2}n^{-1}$ into (72) results in the Anderson-Darling test statistic used for this study

$$A^* = \left(1 + \frac{\sqrt{2}}{n}\right) \cdot \sum_{i}^{n} \frac{1-2i}{n} \left[\ln\left(1 - \exp\left[-\left(\frac{t}{\widehat{\alpha}_{MLE}}\right)^{\widehat{\beta}_{MLE}}\right]\right) - \left(\frac{t_i}{\widehat{\alpha}_{MLE}}\right)^{\widehat{\beta}_{MLE}}\right] - n. \quad (69)$$

For each iteration, the observed significance level (p-value) of the Anderson-Darling test was found using the critical value tables in [107, p. 146]. To demonstrate how the Weibull fit the simulated data generated across the sixty test scenarios, each iteration of the model was considered a Bernoulli trial where success was defined as a test with a p-value equal to or greater than 0.05. Figure 29 displays the results of the analysis. The figure shows that the Anderson-Darling test fails to reject the Weibull for all regions of the design space with the exception of the low-quality, high-complexity and low-complexity, high-quality factor level combinations.

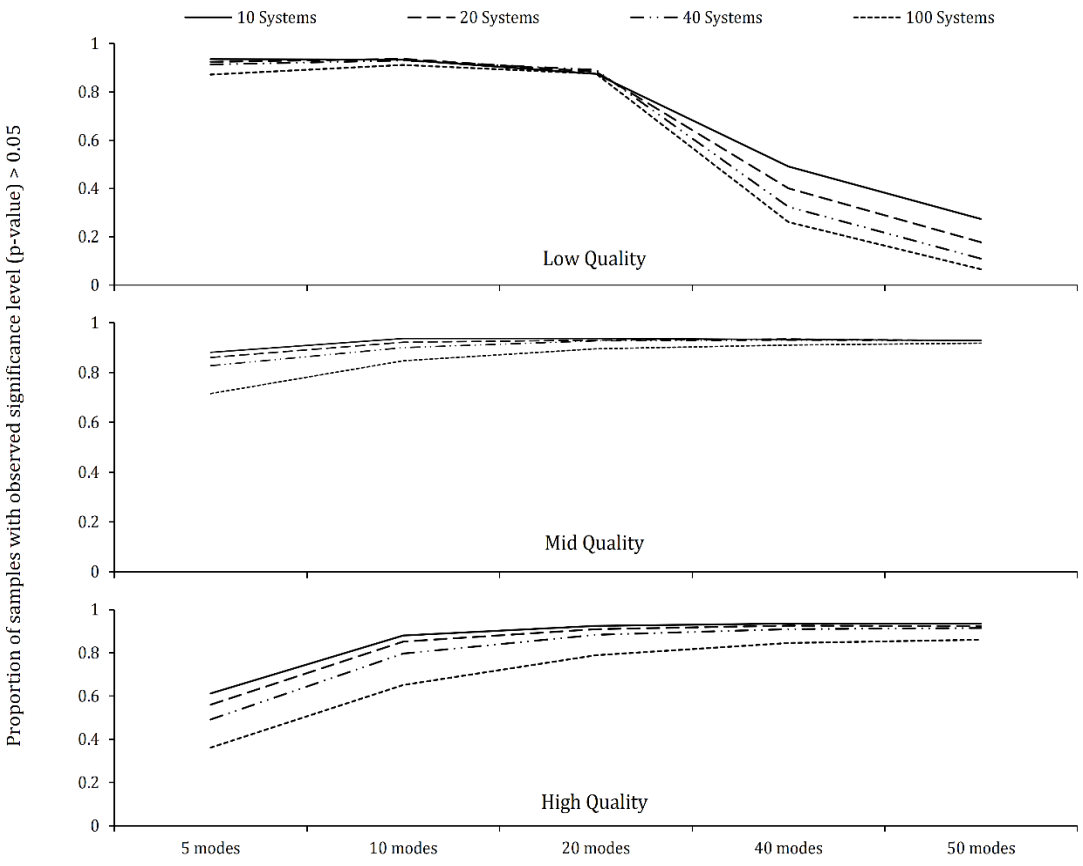

Figure 29 – Results of the Anderson-Darling goodness of fit tests showing the fraction of tests with significance levels greater than 0.05 for each combination of quality level sample size and system complexity

In low-quality, high-complexity systems, many design defects or other potential causes of infant mortality failure are assumed to exist. Thus, the simulated failure observations are concentrated in a time interval where failure occurs almost instantaneously. Fitting data sets such as these with a Weibull distribution implies a shape parameter $\beta \ll 1$. Observing the $AICc$ histograms in Appendix A for systems with these factor level combinations indicates that in a large majority of the simulated data sets the Weibull is preferred model. Combining the results displayed in these histograms with those shown in Figure 29 suggests that the low quality level factor setting may be too low to represent a qualitative accelerated life test. For low-complexity, high-quality systems, histograms of the observations indicate that multimodal density functions result in a significant proportion of the generated data sets. It is in this design region that non-monotone hazard rate functions such as the "bathtub" curve are observed. In practice, qualitative accelerated life tests are typically performed on complex electromechanical devices where a moderate to high level of initial quality is expected.

Observing Figure 29 it is evident that for systems with these characteristics the Weibull distribution is adequate to model an overwhelming majority of the simulated data sets. This result implies that the procedure described above for projecting system reliability after corrective action in the single failure mode case may also be utilized in qualitative testing for systems with multiple mutually independent failure modes where each follows a separate Weibull law.

Computing the system FEF when multiple competing failure modes exist requires taking into account the number of distinct modes exposed during the test. The Weibull parameters associated with each failure mode are unknown as is the contribution each mode makes on the overall failure rate of the system. The system failure rate function, $h(t) = \beta t^{\beta} \alpha^{-\beta}$, represents the sum of the unknown failure rates of each of the embedded modes in the system $h(t) = \sum_j^J h_j(t)$. Assuming each distinct failure mode discovered during the test contributes to the system failure rate equally, the system FEF, P, can be expressed as

$$\mathrm{P} = \frac{\sum_{m=1}^{M} \rho_m}{M} \tag{70}$$

where M denotes the number of distinct failure modes discovered during testing. For the example scenario used to describe the single failure mode case, temperature was the accelerating variable and the Arrhenius model was used to describe relationship between median life and temperature. Expanding this functional relationship for the multiple failure mode case, the projected system reliability after corrective action, with respect to the change in activation energy is expressed as

$$A = \ln\left[\left(1 - \frac{\sum_{m=1}^{M} \rho_m}{M}\right)^{-\frac{T_{use}}{\beta}}\right] + A. \tag{71}$$

And the projected probability density function of system life at T_{use} after corrective action is expressed as

$$P(X \le x; T) = 1 - \exp\left(\frac{x}{C \exp\left[\left(\ln\left[\left(1 - M^{-1} \cdot \sum_{m=1}^{M} \rho_m\right)^{-\frac{T_{use}}{\beta}}\right] + A\right) T_{use}^{-1}\right]}\right)^{\beta}. \tag{72}$$

Conclusions and Future Work

A modified accelerated life test and data analysis procedure was presented. The proposed approach is the first to integrate lifetime acceleration models with reliability growth planning to estimate reliability improvement after implementing corrective action. The procedure was assessed by first considering the case where the systems under test contained a single dominant failure mode. Utilizing the results from a published accelerated life test [18], it was demonstrated that the likelihood of attaining a reliability requirement could be estimated with the proposed procedure. Further, it was shown that by using the proposed procedure, an indication that a system may fail to reach the reliability requirement could be obtained by testing fifty percent less samples than are required by traditional quantitative accelerated life test methods. This result suggests that a significant reduction in the time and cost associated with system testing could be realized as the proposed method may be used to signal a need to reassess a product's design or reallocate testing resources to avoid unnecessary maintenance costs or an expensive redesign after product release.

The procedure was then extended to the case where each system under test contained multiple independent competing failure modes. It was shown that if the time to occurrence distribution for each respective failure mode follows a distinct Weibull law, the observed system failure times follow a poly-Weibull distribution. However, as result of a Monte Carlo simulation it was shown that under certain testing scenarios, specific to qualitative accelerated life testing, the failure data

generated by systems with multiple independent competing failure modes can be modeled with a Weibull distribution. Thus, the proposed model utilizes the Weibull distribution to estimate system reliability after one or more failure modes have been discovered and removed.

VI. Conclusions and Recommendations

Dissertation Summary

Merging reliability growth and accelerated life testing is a relatively new concept driven, in part, by the complexity and widespread use of modern electronic devices. In commercial environments, where several product types may be in development simultaneously, little time may exist to certify that the reliability of a particular product meets the reliability requirement. Qualitative accelerated test methods are therefore often utilized to ensure a product's readiness for market introduction and to avoid costly redesigns caused by critical failures that may not be discovered in traditional reliability testing. While published case-studies attest to the effectiveness of qualitative life testing [15, 16, 14] in improving product reliabilty, the capability to translate the limited failure data into a meaningful measure of reliability improvement does not exist. Thus a goal of this dissertation research was to bridge the gap between quantitative and qualitative accelerated reliability test methods.

In Chapter III, several commonly used accelerated failure data analysis methods were highlighted to show their limitations with respect to estimating system reliability from the data obtained in a simple qualitative accelerated stress test. It was shown that these methods either cannot incorporate design changes resulting from corrective action or over-simplify the reliability growth process to avoid estimating the parameters of the life-stress relationship. It was demonstrated that any corrective action that improves a product's reliability changes, by definition, the

way the product responds to stress and therefore also alters the acceleration factor. Further, the parameters of the stress life relationship must be estimated for each distinct failure mode discovered during the test. Chapter III concluded with a discussion on several potential paths for future research with respect to improving the parameter estimation in follow-on testing.

Because accelerated testing models correspond to individual failure modes, the proper approach for modeling reliability growth of a complex system in an accelerated test scenario requires an accounting of the failure rate changes for each mode separately and integrating these changes into a single measure of system reliability. For each failure event, only one risk can be attributed as its cause, where a risk may be defined as either a single failure mode or the combination of multiple modes acting simultaneously. It was shown in Chapter IV that if it can be assumed that the time to occurrence distribution for each competing mode follows a Weibull distribution with mode specific parameters α_j and β_j, $j = 1, \ldots, J$, the system lifetime can be modeled with the poly-Weibull distribution with vector valued parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}$. The poly-Weibull was then presented as an alternative to the Weibull distribution for fitting data sets where the hazard function is monotone, non-monotone or even bathtub shaped. A numerical and analytical procedure for obtaining the poly-Weibull maximum likelihood parameter estimates and asymptotic standard errors was presented which allows reliability engineers to draw conclusions about system lifetime with a stated degree of confidence.

It was also shown that the poly-Weibull distribution is capable of fitting reference data sets known to have bathtub-shaped hazard rate functions better than the best modified Weibull models in the literature. The goodness of fits for two forms of the poly-Weibull distribution, the bi-Weibull and tri-Weibull, were assessed using the Akaike information criterion and Kolmogorov-Smirnov test statistic. The results showed that the bi-Weibull and the tri-Weibull outperform other modified Weibull distributions with respect to their fit of data for which the hazard rate function is bathtub shaped.

Lastly, a model was presented in Chapter V to characterize a product's likelihood of attaining a reliability requirement after implementing corrective action to remove one or more embedded failure modes. It was that shown under the assumptions of qualitative accelerated life testing the Weibull distribution may be used to model accelerated life test data when multiple independent failure modes are observed. An example test scenario [18] was discussed in which a reliability enhancement program was abandoned after multiple tests showed that the new design would not meet the required median life. Our results indicated that had our proposed model been used to estimate the reliability risk, a significant cost and time savings could have been realized by using less than half of the samples required for the original test. Further, in Chapter V it was shown that the projection model can be used to conduct trade off analyses as a basis for reviewing the reliability requirements. These are important contributions as the model allows reliability engineers to design better systems at less cost and in less time than would be possible otherwise.

Conclusions and Future Work

Qualitative accelerated life test methods were initially developed with the intent of discovering and removing previously unknown failure modes from early system prototypes. As opposed to quantitative accelerated life testing, these methods are not designed with life estimation or prediction in mind. In a few of these tests, such as highly accelerated stress testing, extensive testing and data analysis has been presented in the literature to derive a life-stress relationship for specific combinations of stressors. In general, however, the ability to devise ever-harsher test environments to discover more failure modes in less time will always outpace the rate at which mathematically sound tools can be generated to estimate or predict product reliability.

The goal of this dissertation was to provide first step toward bridging the gap separating quantitative and qualitative accelerated life testing. The most obvious next steps to be taken in further research to improve the reliability growth projection model involve extending the test scenario to multiple stressors and to include multiple corrective action periods to remove failure modes discovered in subsequent tests. Both of these paths are discussed in the sections below.

Extending the Model to Multiple Corrective Action Periods

For systems with multiple failure modes, it is unlikely that a single test with a single corrective action period can uncover all of the failure modes that could occur in the use environment. In reality, many corrective action periods may be required. Extending the projection model to testing scenarios with multiple corrective action

periods requires that the corrective actions to remove the failure modes discovered during the preceding tests are not cost prohibitive. The reliability growth attained through corrective action in an accelerated test depends on the likelihood of the failure actually occurring in operation [54, 79, 17]. For the single failure mode case, it was shown in Chapter V that each corrective action increases the activation energy by $A' - A$. With each successive improvement, the energy required to activate a failure mode approaches the energy required for a non-representative failure to occur. Nelson [54] describes an example in which a costly program was initiated to remove a failure mode that, it was later determined, would have never occurred in the use environment. Thus, the end result of the program was a substantial amount of wasted resources and no actual reliability improvement.

One possible way to address this is by introducing an adjusted FEF of the form

$$\rho^* = \rho \cdot d(S) \tag{73}$$

where ρ is the estimated FEF value for a given failure mode and $d(S) \in (0,1)$ is a multiplicative factor that scales ρ based on the level of stress S at which the test is performed. When the stress applied to the system is at or below the field-use stress, $S \leq S_{use}$, $d(S)$ is defined to be equal to unity. Conversely, when the applied stress is above some limit stress $S \geq S_{limit}$,where failure occurs almost immediately due to an unrecoverable failure, $d(S) \equiv 0$. Examples of unrecoverable failures include the melting or phase change of a material as result of elevated temperatures, catastrophic mechanical failure due to elevated vibrational stresses and dielectric

breakdown due to voltage overstress. For $S_{use} < S < S_{limit}$, a possible form of $d(S)$ is a monotone decreasing sigmoid function such as

$$d(S) = 1 - \frac{1}{1 + \exp\left[-\left(\frac{w}{S_{limit} - S_{use}}\right)\left(S - \frac{(S_{use} + S_{limit})}{2}\right)\right]} \qquad (74)$$

where the functional form of (74) is derived by modifying the logistic function. The parameter w may be used to represent how quickly $d(S)$ approaches zero for a given system and testing scenario. Figure 30 illustrates the shape of $d(S)$ for various values of S_{limit} and for $w = 10$. The true form of $d(S)$ is unknown, however the figure illustrates that accelerated reliability growth test implies a trade-off between the rate at which failures occur and the level of improvement made. By introducing an adjusted FEF value to the model developed in Chapter V, test planners can balance the level of stress applied with the expected reliability improvement.

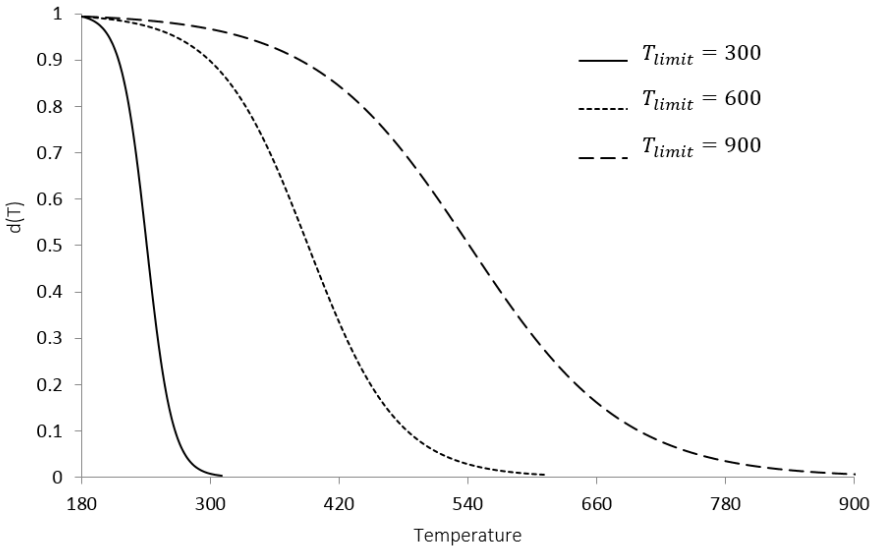

Figure 30 – Sigmoid shape of $d(S)$

Extending the Model to Multiple Stressor Tests

Extending the reliability growth projection model developed in Chapter V to address stress states in which more than one stressor is used would require estimating the parameters of the life-stress relationship for each stress separately along with any interaction terms that may be significant. Even, if it can be assumed that no interaction exists between the stressors (a very restrictive assumption) each stressor added to a design would bring two additional parameters that must be estimated. Estimating these parameters would require multiple observations of each failure mode from tests performed at distinct design points. By dividing an already small test sample across a greater number of design points would likely result in parameter estimates of the life-stress relationship that have large variances. This would lead to confidence intervals that would be far too wide to be meaningful. Unfortunately, financial realities limit the number of units that can be made available for early system testing and therefore constrain the decision space as to feasible testing options. It is our hope that in spite of these constraints progress can be made to further bridge the gap between qualitative and quantitative accelerated life testing.

Bibliography

[1] M. Donley, "Air Force Modernization and Recapitalization Strategy," Highbeam Research, 18 November 2008. [Online]. Available: http://www.highbeam.com/doc/1G1-190285075.html. [Accessed October 2012].

[2] R. Gates, "Science and Technology (S&T) Priorities for Fiscal Years 2013-17 Planning," 19 April 2011. [Online]. Available: http://www.acq.osd.mil/chieftechnologist/publications/docs/OSD%2002073-11.pdf. [Accessed June 2012].

[3] R. Neches, "Engineered Resilient Systems (ERS) S&T Priority, Description and Roadmap," 20 December 2011. [Online]. Available: http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA554975. [Accessed June 2012].

[4] M. Gilmore, "Key Issues in Reliability Growth," in *Panel on the Theory and Application of Reliability Growth Modeling in Defense Systems*, Washington, DC, 2011.

[5] Committee On National Statistics, Statistics, Testing and Defense Acquisition: New Approaches and Methodological Improvements, M. Cohen, J. E. Rolph and D. L. Steffey, Eds., Atlanta, GA: The National Academy Press, 1998.

[6] Committee on National Statistics, Reliability Issues for DOD Systems: Report of a Workshop, F. Samaniego and M. Cohen, Eds., Atlanta, GA: The National Academies Press, 2002.

[7] US Department of Defense, Undersecretary for Acquisition, Technology and Logistics, "Reliability Analysis, Planning, Tracking and Reporting," Government Printing Office, Washington, DC, 21 Mar 2011.

[8] GEIA, G-47 Committee - Quality & Reliability Engineering, "Reliability program standard for systems design, development and manufacturing," ANSI Publishing, Washington, DC, 2008.

[9] MIL-HDBK-189C, "Reliability Growth Management," Naval Publications and Forms Center, Philadelphia, PA, 2011.

[10] J. B. Hall, "Implementation Guide for U.S. Army Reliability Policy - A Technical Note," United States Army Evaluation Center, Aberdeen Proving Grounds, MA, 2009.

[11] A. Bhote and K. Bhote, World Class Reliability: Using Multiple Environment Overstress Tests to Make it Happen, Saranac Lake, NY: AMACOM Publishing, 2004.

[12] H. W. McLean, HALT, HASS and HASA Explained, Milwaukee, WI: Quality Press, 2009.

[13] A. J. Porter, "Failure Mode Verification: Applying Highly Accelerated Life Testing to Mechanical Systems," in *International Body Engineering Conference and Exposition*, Detroit, MI, 1998.

[14] R. H. Gusciora, "Why continue to try proving that HALT really works for electronic parts?," in *National Electronics Packaging and Production Conference - Proceedings of the Technical Program*, Des Plaines, IL, 2000.

[15] E. L. Kyser and C. Ascarrunz, "HALT and ESS for quick-to-market scenarios," in *National Electronics Packaging and Production Conference - Proceedings of the Technical Program*, Des Plaines, IL, 1998.

[16] M. Silverman, "Summary of HALT and HASS results at an accelerated reliability test center," in *National Electronics Packaging and Production Conference - Proceedings of the Technical Program*, Des Plaines, IL, 1997.

[17] W. Q. Meeker, "More pitfalls of accelerated tests," *Journal of Quality Technology,* vol. 45, no. 3, pp. 213-222, 2013.

[18] W. Nelson, "Accelerated Life Testing - step-stress models and data analyses," *IEEE Transactions on Reliability,* vol. 29, pp. 103-108, 1980.

[19] G. Bhattacharyya and Z. Soejoeti, "A tampered failure rate model for step-stress accelerated life tests," *Communications in Statistics, Theory and Models,* vol. 18, no. 5, pp. 1627-1643, 1989.

[20] T. A. Gooley, W. Leisenring, J. Crowley and B. E. Storer, "Estimation of failure probabilities in the presence of competing risks: new representations of old estimators," *Statistics in Medicine,* vol. 18, no. 6, pp. 695-706, 1999.

[21] H. A. David and M. L. Moeschberger, The theory of competing risks, London, England: Macmillan Publishing, 1978, p. 103.

[22] H. Putter, M. Fiocco and R. Geskus, "Tutorial in Biostatistics: Competing risks and multi-state models," *Statistics in Medicine,* vol. 26, no. 11, pp. 2389-2430, 2007.

[23] F. Siannis, J. Copas and L. Guobing, "Sensitivity analysis for informative censoring in parametric survival models," *Biostatistics,* vol. 6, no. 1, pp. 77-91, 2005.

[24] M. J. Crowder, Classical Competing Risks, New York, NY: Chapman & Hall, 2001, p. 200.

[25] D. R. Cox, "Regression models and life tables (with discussion)," *Journal of the Royal Statistical Society, Series B (Methodological),* vol. 34, no. 2, pp. 187-220, 1972.

[26] Z. Ma and A. W. Krings, "Competing risks analysis of reliability, survivability and prognostics and health management (PHM)," in *IEEE Aerospace Conference*, Big Sky, MT, 2008.

[27] D. R. Cox, "The analysis of exponentially distributed life-times with two types of failure," *Journal of the Royal Statistical Society, Series B (Methodological),* vol. 21, no. 2, pp. 411-421, 1959.

[28] A. Tsiatis, "A nonidentifiability aspect of the problem of competing risks," *Proceedings of the National Academy of Science,* vol. 72, no. 1, pp. 20-22, 1975.

[29] R. J. Caplan, T. F. Pajak and J. D. Cox, "Analysis of the probability and risk of cause specific failure," *International Journal of Radiology Oncology, Biology, Physics,* vol. 29, no. 5, pp. 1183-1186, 1994.

[30] E. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *Journal of the American Statistical Association,* vol. 53, no. 282, pp. 457-481, 1958.

[31] S. Cheng, J. P. Fine and L. J. Wei, "Prediction of cumulative incidence function under the proportional hazards model," *Biometrics,* vol. 54, no. 1, pp. 219-228, 1998.

[32] J. D. Kalbfleisch and R. L. Prentice, The Statistical Analysis of Failure Time Data, Hoboken, NJ: John Wiley & Sons, 1980.

[33] W. J. Corcoran, H. Weingarten and P. W. Zehna, "Estimating reliability after corrective action," *Management Science,* vol. 10, no. 4, pp. 786-795, 1964.

[34] R. C. Dahiya, "Estimation of reliability after corrective action," *IEEE Transactions on Reliability,* Vols. R-26, no. 5, pp. 348-351, 1977.

[35] D. E. Olsen, "Estimating reliability growth," *IEEE Transactions on Reliability,* Vols. R-26, no. 1, pp. 50-53, 1977.

[36] MIL-HDBK-338B, "Electronic Reliability Design Handbook," DoD Reliability Analysis Center, Rome, NY, 1992.

[37] C. E. Ebeling, An Introduction to Reliability and Maintainability Engineering, 2nd ed., Long Grove, IL: Waveland Press, 2009.

[38] J. T. Duane, "Learning curve approach to reliability monitoring," *IEEE Transactions on Aerospace,* vol. 2, no. 2, pp. 563-566, 1964.

[39] L. Crow, "Reliability analysis for complex repairable systems," in *Reliability and Biometry*, Philadelphia, PA: SIAM, 1974, pp. 379-410.

[40] A. Sen, "Analysis of Repairable Systems--Past Present and Future," in *Frontiers in Reliability*, World Scientific, 1998, pp. 317-336.

[41] J. B. Hall, "Dissertation - Methodology For Evaluating Reliability Growth Programs of Discrete Systems," University of Maryland, College Park, MD, 2008.

[42] W. Q. Meeker and M. Hamada, "Statistical tools for the rapid development and evaluation of high-reliability products," *IEEE Transactions on Reliability,* vol. 44, no. 2, pp. 187-198, 1995.

[43] D. Kececioglu and F. B. Sun, Environmental Stress Screening: Its Quantification Optimization and Management, DEStech Publishing: Lancaster, PA, 2002.

[44] G. K. Hobbs, Accelerated Reliability Engineering, Hoboken, NJ: John Wiley & Sons, 2000.

[45] L. A. Escobar and W. Q. Meeker, "A review of accelerated test models," *Statistical Science,* vol. 21, no. 4, pp. 552-577, 2008.

[46] M. Silverman, "Why HALT Cannot Produce a Meaningful MTBF Number and Why This Should Not be a Concern," [Online]. Available: http://www.opsalacarte.com/pdfs/Tech_Papers/MTBF_Prediction_HALT.pdf. [Accessed Jun 2012].

[47] L. Edson, *Calibrated Accelerated Life Testing - (General Motors Test Procedure GMW8758),* General Motors Worldwide, 2011.

[48] M. K. Brand and H. W. McLean, "Method for estimating changes in product life for a redesigned product". United States Patent 7,149,673, 2006.

[49] M. K. Brand and H. W. McLean, "Method for estimating changes in product life resulting from HALT using exponential acceleration". United States Patent 7,120,566, 2007.

[50] M. K. Brand and H. W. McLean, "Method for estimating changes in product life resulting from HALT using quadratic acceleration". United States Patent 7,260,509, 2007.

[51] H. W. McLean, "A method of estimating product field failure rate from the results of HALT," in *IEEE Accelerated Stress Testing and Reliability Workshop*, Portland, OR, 2008.

[52] D. Bai, "Optimum simple step-stress accelerated life tests with censoring," *IEEE Transactions on Reliability,* vol. 38, no. 5, pp. 528-532, Dec 1989.

[53] C.-M. Laio, "Optimal design for step stress degradation testing," *IEEE Transactions on Reliability,* vol. 55, no. 1, pp. 59-66, 2006.

[54] W. B. Nelson, Accelerated Testing: statistical models, test plans and data analyses, Hoboken, NJ: John Wiley & Sons, 1980.

[55] D. Kececioglu, Accelerated life testing and reliability determination lecture notes, Tuscon, AZ: The University of Arizona - Microcampus , 1986.

[56] W. Zhao, Y. Tang and E. Elsayed, "A general accelerated life model for step stress testing," *IIE Transactions,* vol. 37, no. 11, pp. 1059-1069, 2005.

[57] D. Nicholls and P. Lein, Achieving System Reliability Growth Through Robust Design and Test, Rome, NY: Reliability Information and Analysis Center, 2011.

[58] P. O'Connor and A. Kleyner, Practical Reliability Engineering, 5th ed., Hoboken, NJ: John Wiley & Sons, 2011.

[59] H. Pham, Springer Handbook of Engineering Statistics, New York, NY: Springer, 2006.

[60] L. A. Escobar and W. Q. Meeker, "Planning Accelerated Life Tests with Two or More Experimental Factors," *Technometrics,* vol. 37, no. 4, pp. 411-427, 1995.

[61] C. Huber-Carol, N. Balakrishnan, M. Nikulin and M. Mesbah, Goodness of Fit Tests and Model Validity, Boston, MA: Birkhauser, 2002.

[62] L. J. Freeman and G. G. Vining, "Reliability data analysis for life test designed experiments with subsampling," *Quality and Reliability Engineering International,* vol. 29, no. 4, pp. 509-519, 2013.

[63] M. Zelen, "Factorial experiments in life testing," *Technometrics,* vol. 1, no. 3, pp. 269-288, 1959.

[64] W. Weibull, Fatigue testing and analysis of results, New York, NY: Pergamon Press, 1961.

[65] M. H. Degroot and P. K. Goel, "Bayesian estimation and optimal designs in partially accelerated life testing," *Naval Research Logistics Quarterly,* vol. 26, no. 2, pp. 223-235, Jun 1979.

[66] Joint Electron Device Engineering Council, "Reliability Qualification of Semiconductor Devices Based on Physics of Failure Risk and Opportunity Assessment," JEDEC, 2008.

[67] G. Drake, "Engineering Design Analysis (Physics of Failure)," in *Proceedings of the Reliability and Maintainability Symposium*, San Jose, CA, 2010.

[68] M. Modarres, M. Kaminskiy and V. Krivtsov, Reliability Engineering and Risk Analysis. A Practical Guide, New York, NY: Marcel Dekker, 1999.

[69] R. Heine, "The use of physics of failure for reliability improvement and addressing modularity issues in evaluation and physical testing," in *National Defense Industrial Association - Physics Based Modeling in Design and Development for U.S. Defense*, Denver, CO, 2011.

[70] M. Pecht, A. Dasgupta, D. Barker and C. Leonard, "The reliability physics approach to failure prediction modeling," *Quality and Reliability Engineering International,* vol. 6, no. 4, pp. 267-273, 1990.

[71] Rome Air Defense Center, "Mil-HDBK-217 Revision F Notice 2," DoD Publishing, Washington DC, 1995.

[72] M. Cushing, D. Mortin, T. Stadterman and A. Malhotra, "Comparison of Electronics-Reliability Assessment Approaches," *IEEE Transactions on Reliability,* vol. 42, no. 4, pp. 542-546, 1993.

[73] M. Mendel, "The case for probabilistic physics of failure," *Reliability and Maintenance of Complex Systems,* vol. 154, pp. 70-82, 1996.

[74] I. Snook, J. M. Marshall and R. M. Newman, "Physics of failure as an integrated part of design for reliability," in *Proceedings of the Reliability and Maintainability Symposium*, Tampa, FL, 2003.

[75] H. Qi, "A rapid life-prediction approach for PGBA solder joints under combined thermal cycling and vibration locading conditions," *IEEE*

Transactions on Components and Packing Technologies, vol. 32, no. 2, pp. 283-292, 2009.

[76] IPC - Association Connecting Electronic Industries, "IPC-9701A - Performance Test Methods and Qualification Requirements for Surface Mount Solder Attachments," IPC, 2006.

[77] A. A. Feinberg, "Accelerated reliability growth models," *Journal of the Institute of Environmental Sciences,* pp. 17-23, 1994.

[78] R. Strunz and J. W. Herrmann, "Planning, Tracking and Projecting Reliability Growth: A Bayesian Approach," in *Proceedings of the Reliability and Maintainability Symposium*, Orlando, FL, 2012.

[79] W. Q. Meeker and L. A. Escobar, "Pitfalls of accelerated life testing," *IEEE Transactions on Reliability,* vol. 47, no. 2, pp. 114-118, 1998.

[80] H. McLean, "From HALT results to an accurate field MTBF estimate," in *Proceedings of the Reliability and Maintainability Symposium*, San Jose, CA, 2010.

[81] Semiconductor Equipment and Materials International, "SEMI E10/0312 - Specification for Definition and Measurement of Equipment Reliability, Availability, and Maintainability (RAM) and Utilization," Semiconductor Equipment and Materials International, 2012.

[82] L. Gao, W. Chen, J. Liu, J. Pan and P. Qian, "Design Criteria for Planning Multiple Stresses Accelerated Life Test," in *International Conference of Reliability, Maintainability and Safety*, Guiyang, China, 2001.

[83] M. Aarset, "How to identify a bathtub hazard rate," *IEEE Transactions of Reliability,* vol. 36, no. 1, pp. 106-108, 1987.

[84] W. Q. Meeker and L. A. Escobar, Statistical Methods for Reliability Data, Hoboken, NJ: John Wiley & Sons, 1998.

[85] J. Berger and D. Sun, "Bayesian analysis for the Poly-Weibull distribution," *Journal of the American Statistical Association,* vol. 88, no. 424, pp. 1412-1418, 1993.

[86] S. J. Almalki and J. Yuan, "A new modified Weibull distribution," *Reliability Engineering and System Safety,* vol. 111, pp. 164-170, 2013.

[87] A. M. Sarhan and J. Apaloo, "Exponentiated modified Weibull extension distribution," *Reliability Engineering and System Safety,* vol. 112, pp. 137-144, 2013.

[88] C. Lai, M. Xie and D. Murthy, "A modified Weibull Distribution," *IEEE Transactions on Reliability,* vol. 52, no. 1, pp. 33-37, 2003.

[89] M. Xie and C. Lai, "Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function," *Reliability Engineering and System Safety,* vol. 52, no. 1, pp. 87-93, 1995.

[90] A. M. Sarhan and M. Zaindin, "Modified Weibull Distribution," *Applied Sciences,* vol. 11, pp. 123-136, 2009.

[91] G. O. Silva, E. M. Ortega and G. M. Cordiero, "The beta modified Weibull distribution," *Lifetime Data Analysis,* vol. 16, no. 3, pp. 409-430, 2010.

[92] G. Mudholkar and D. Srivastava, "Exponentiated Weibull family for analyzing bathtub failure rate data," *IEEE Transactions on Reliability,* vol. 42, no. 2, pp. 233-302, 1993.

[93] A. El-Gohary, A. Alshamrani and A. Al-Otaibi, "The Generalized Gompertz Distribution," *Applied Mathematical Modeling,* vol. 37, no. 1, pp. 13-24, 2013.

[94] M. Xie, Y. Tang and T. Goh, "A modified Weibull extension with bathtub-shaped failure rate function," *Reliability Engineering and System Safety,* vol. 76, no. 3, pp. 279-285, 2002.

[95] A. C. Davison and F. Louzada-Neto, "Inference for the poly-Weibull model," *Journal of the Royal Statistical Society, Series D (The Statistician),* vol. 49, no. 2, pp. 189-196, 2000.

[96] L. Tierney and J. Kadane, "Accurate approximations for posterior moments and marginal densities," *Journal of the Americal Statistical Association,* vol. 81, no. 393, pp. 82-86, 1986.

[97] A. Smith and A. E. Gelfand, "Bayesian statistics without tears: a sampling-resampling perspective," *American Statistian,* vol. 46, no. 2, pp. 84-88, 1992.

[98] R. Byrd, P. Lu, J. Nocedal and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing,* vol. 16, no. 5, pp. 1190-1208, 1995.

[99] J. F. Lawless, Statistical Models and Methods for Lifetime Data, Hoboken, NJ: John Wiley & Sons Inc., 1982.

[100] J. Seo, M. Jung and C. Kim, "Design of accelerated life test sampling plans with a non-constant shape parameter," *European Journal of Operations Research,* vol. 197, no. 2, pp. 659-666, 2009.

[101] K. Hamed and A. R. Rao, Flood Frequency Analysis, Boca Raton, Florida: CRC Press, 2010.

[102] T. Jin, Z. Xiong and P. Wand, "Minimize system failure rate considering variations of electronics components lifetime data," in *International Conference of Asian Green Electronics*, Shanghai, China, 2005.

[103] D. Wackerly, W. Mendenhall and R. Scheaffer, Mathematical Statistics with Applications, Seventh Edition, Belmont, CA: Thomson Higher Education, 2008.

[104] C. M. Hurvich and C. L. Tsai, "Regression and time series model selection in small sample sizes," *Biometrika,* vol. 76, no. 2, pp. 297-307, 1989.

[105] T. Anderson and D. A. Darling, "Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes," *Annals of Mathematical Statistics,* vol. 23, no. 2, pp. 193-212, 1952.

[106] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *Journal of the American Statistical Association,* vol. 69, no. 347, pp. 730-737, 1974.

[107] R. B. D'Agostino and M. A. Stephens, Goodness of Fit Techniques, New York, NY: Marcel Dekker, Inc., 1986.

Appendix A: Poly–Weibull Observed Fisher Information Equations

The observed Fisher information for the bi-Weibull and tri-Weibull distributions are expressed respectively as

$$\mathcal{I}_{BW}(\boldsymbol{\alpha},\boldsymbol{\beta}) = -\nabla\nabla^{\mathrm{T}}\mathcal{L}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \begin{bmatrix} \frac{\partial^2\mathcal{L}}{\partial\beta_1^2} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\alpha_2} \\ \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_2^2} & \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\alpha_2} \\ \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1^2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\alpha_2} \\ \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2^2} \end{bmatrix}$$

and

$$\mathcal{I}_{TW}(\boldsymbol{\alpha},\boldsymbol{\beta}) = -\nabla\nabla^{\mathrm{T}}\mathcal{L}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \begin{bmatrix} \frac{\partial^2\mathcal{L}}{\partial\beta_1^2} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\beta_3} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\alpha_2} & \frac{\partial^2\mathcal{L}}{\partial\beta_1\partial\alpha_3} \\ \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_2^2} & \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\beta_3} & \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\alpha_2} & \frac{\partial^2\mathcal{L}}{\partial\beta_2\partial\alpha_3} \\ \frac{\partial^2\mathcal{L}}{\partial\beta_3\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_3\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\beta_3^2} & \frac{\partial^2\mathcal{L}}{\partial\beta_3\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\beta_3\partial\alpha_2} & \frac{\partial^2\mathcal{L}}{\partial\beta_3\partial\alpha_3} \\ \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\beta_3} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1^2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\alpha_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_1\partial\alpha_3} \\ \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\beta_3} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2^2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_2\partial\alpha_3} \\ \frac{\partial^2\mathcal{L}}{\partial\alpha_3\partial\beta_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_3\partial\beta_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_3\partial\beta_3} & \frac{\partial^2\mathcal{L}}{\partial\alpha_3\partial\alpha_1} & \frac{\partial^2\mathcal{L}}{\partial\alpha_3\partial\alpha_2} & \frac{\partial^2\mathcal{L}}{\partial\alpha_3^2} \end{bmatrix}.$$

The elements of these matrices are listed below, where $j, k = 1, 2, 3, \ldots$

$$\frac{\partial\mathcal{L}}{\partial\beta_j} = \sum_i^N h_j(t_i)\left[\frac{\left(\log\left(t_i/\alpha_j\right)+\beta_j^{-1}\right)}{h(t_i)} - \left(\frac{t_i}{\beta_j}\right)\log\left(\frac{t_i}{\alpha_j}\right)\right]$$

$$\frac{\partial^2\mathcal{L}}{\partial\beta_j^2} = \sum_i^N\left[\frac{h_j(t_i)\cdot\log\left(\alpha_j/t_i\right)\left(\log\left(\alpha_j/t_i\right)-2\beta_j^{-1}\right)\cdot h(t_i) - \left(h_j(t_i)\left(\log\left(t_i/\alpha_j\right)+\beta_j^{-1}\right)\right)^2}{h(t_i)^2} - \left(\frac{t_i}{\alpha_j}\right)^{\beta_j}\log^2\left(\frac{t_i}{\alpha_j}\right)\right]$$

$$\frac{\partial^2\mathcal{L}}{\partial\beta_j\partial\beta_k} = \sum_i^N -\frac{h_j(t_i)\left(\log\left(t_i/\alpha_j\right)+\beta_j^{-1}\right)\cdot h_k(t_i)\left(\log\left(t_i/\alpha_k\right)+\beta_k^{-1}\right)}{h(t_i)^2}$$

$$\frac{\partial^2\mathcal{L}}{\partial\beta_j\partial\alpha_j} = \sum_i^N\frac{\beta_j}{\alpha_j}\left[\frac{\left(-h(t_i)h_j(t_i)\left(\log\left(t_i/\alpha_j\right)+2\beta_j^{-1}\right)+h_j(t_i)^2\left(\log\left(t_i/\alpha_j\right)+\beta_j^{-1}\right)\right)}{h(t_i)^2} + \left(\frac{t_i}{\alpha_j}\right)^{\beta_j}\left(\log\left(\frac{t_i}{\alpha}\right)+\beta_j^{-1}\right)\right]$$

$$\frac{\partial^2 \mathcal{L}}{\partial \beta_j \partial \alpha_k} = \sum_i^N h_j(t_i) h_k(t_i) \frac{\beta_j}{\alpha_j} \left[\frac{\left(\log\left(t_i/\alpha_j\right) + \beta_j^{-1} \right)}{h(t_i)^2} \right]$$

$$\frac{\partial \mathcal{L}}{\partial \alpha_j} = \sum_i^N h_j(t_i) \frac{\beta_j}{\alpha_j} \left[\frac{t_i}{\beta_j} - \frac{1}{h(t_i)} \right]$$

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha_j^2} = \sum_i^N \frac{\beta_j}{\alpha_j} \frac{\left[h_j(t_i) h(t_i) \left(\beta_j + 1\right) - \left(\beta_j h_j(t_i) \right)^2 \right]}{h(t_i)^2} - (\beta + 1) \left(\frac{t_i}{\alpha_j} \right)^{\beta_j}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha_j \partial \alpha_k} = \sum_i^N - \frac{\frac{\beta_j}{\alpha_j} h_j(t_i) \cdot \frac{\beta_k}{\alpha_k} h_k(t_i)}{h(t_i)^2}$$

Appendix B: Histograms of Corrected Akaike Information Criterion Values ($AICc$) for the Weibull, Bi-Weibull and Tri-Weibull Distributions

Using the n observations in each of the simulated samples described in Chapter V, the relative performance of the Weibull distribution was compared to the bi- and tri-Weibull using the Akaike Information Criterion corrected for sample size ($AICc$) [104]. The corrected Akaike Information Criterion is

$$AICc = 2k - 2(\pounds) + \frac{2k(k+1)}{n-k-1} \tag{75}$$

where k denotes the number of parameters in each candidate model and $\pounds$ is the value of the log-likelihood function evaluated at the maximum likelihood parameter estimates. For each simulated data set, the $AICc$ of one of the three candidate distributions will have the lowest value. The model corresponding to $AICc_{min}$ has the highest relative likelihood of minimizing the information loss. This likelihood is

$$L(m_i|data) = \exp[0.5(AICc_{min} - AICc_i)] \in [0,1]. \tag{76}$$

This appendix displays histograms of the likelihood values obtained for the Weibull, bi-Weibull and tri-Weibull distributions for each of the sixty testing scenarios. In each figure the histogram for the Weibull distribution is presented in the top plot while the bi-Weibull and tri-Weibull are shown in the middle and bottom plots, respectively. The figures are organized, first according to the number of embedded failure modes, then by the number of systems in each sample and finally by the quality level of each system.

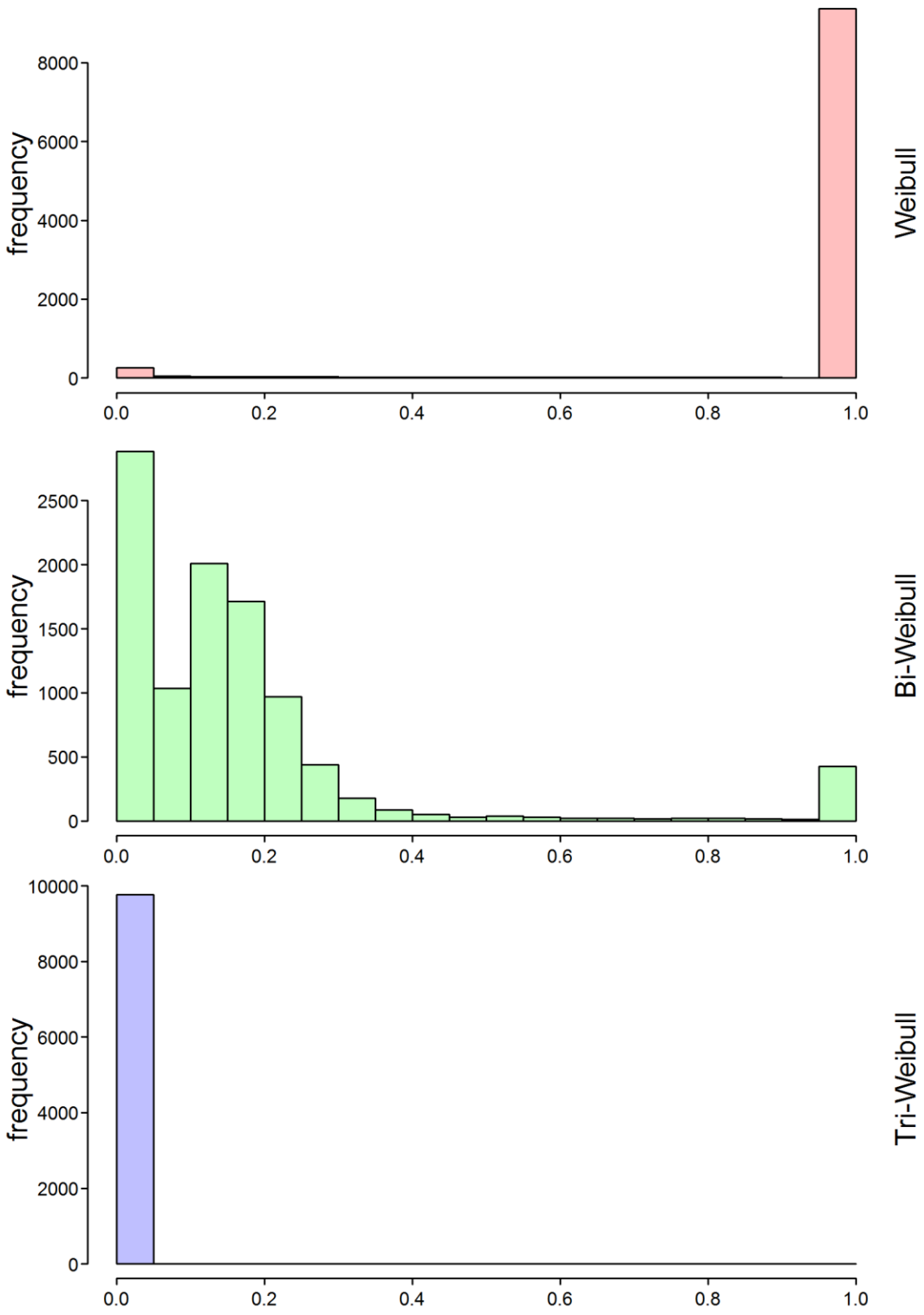

Figure 31 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 10$ systems, quality = low)

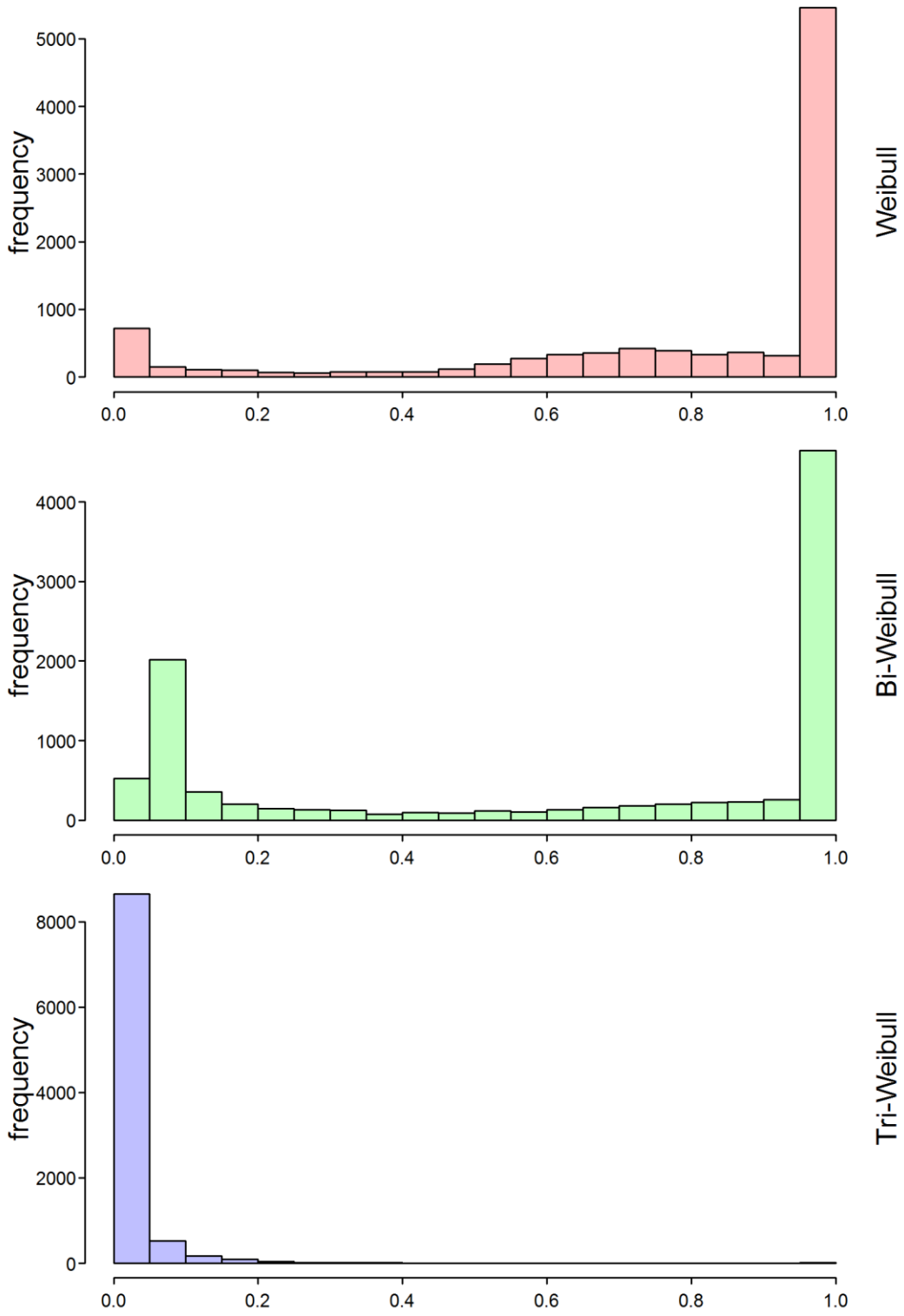

Figure 32 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 20$ systems, quality = low)

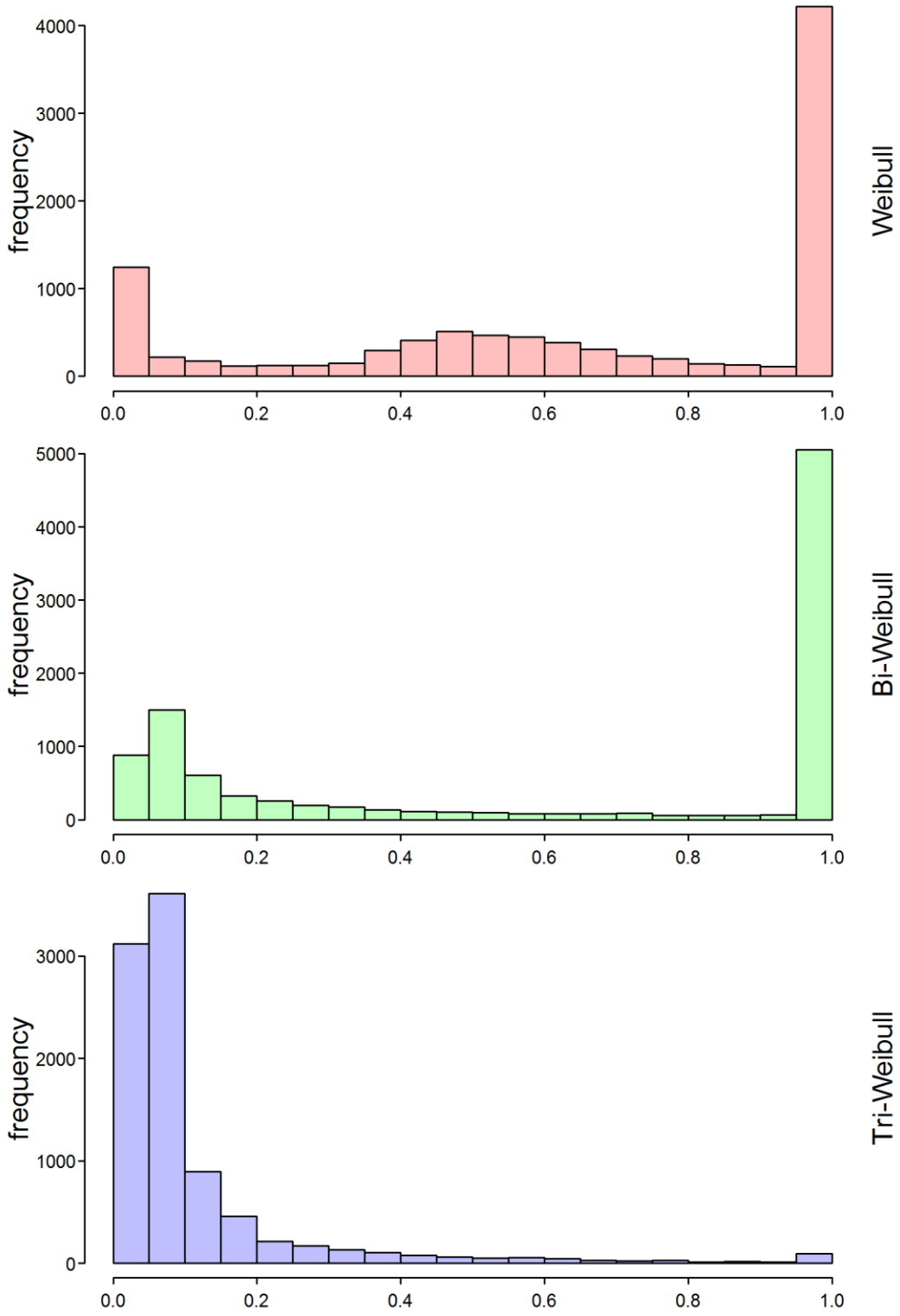

Figure 33 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 40$ systems, quality = low)

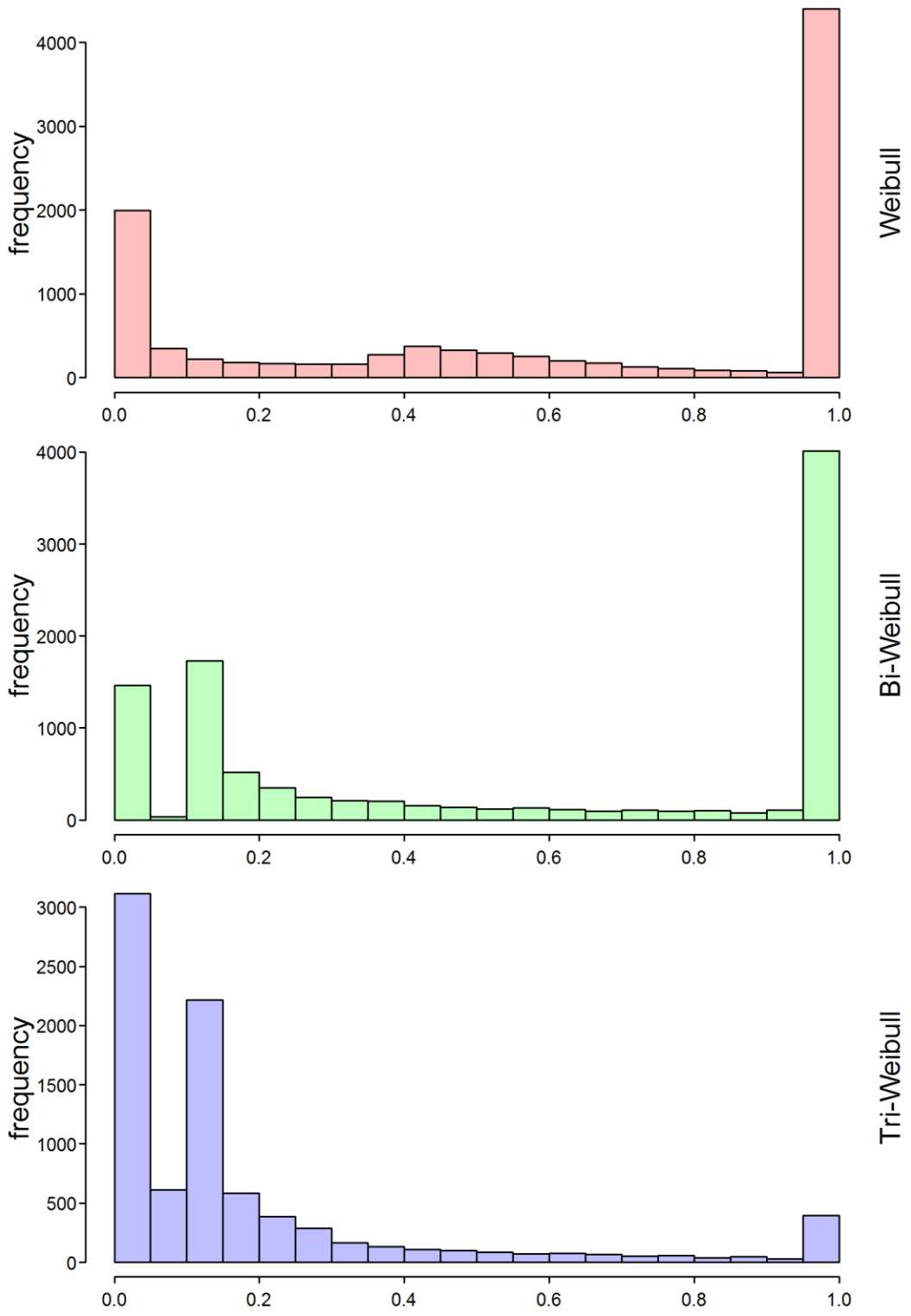

Figure 34 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 100$ systems, quality = low)

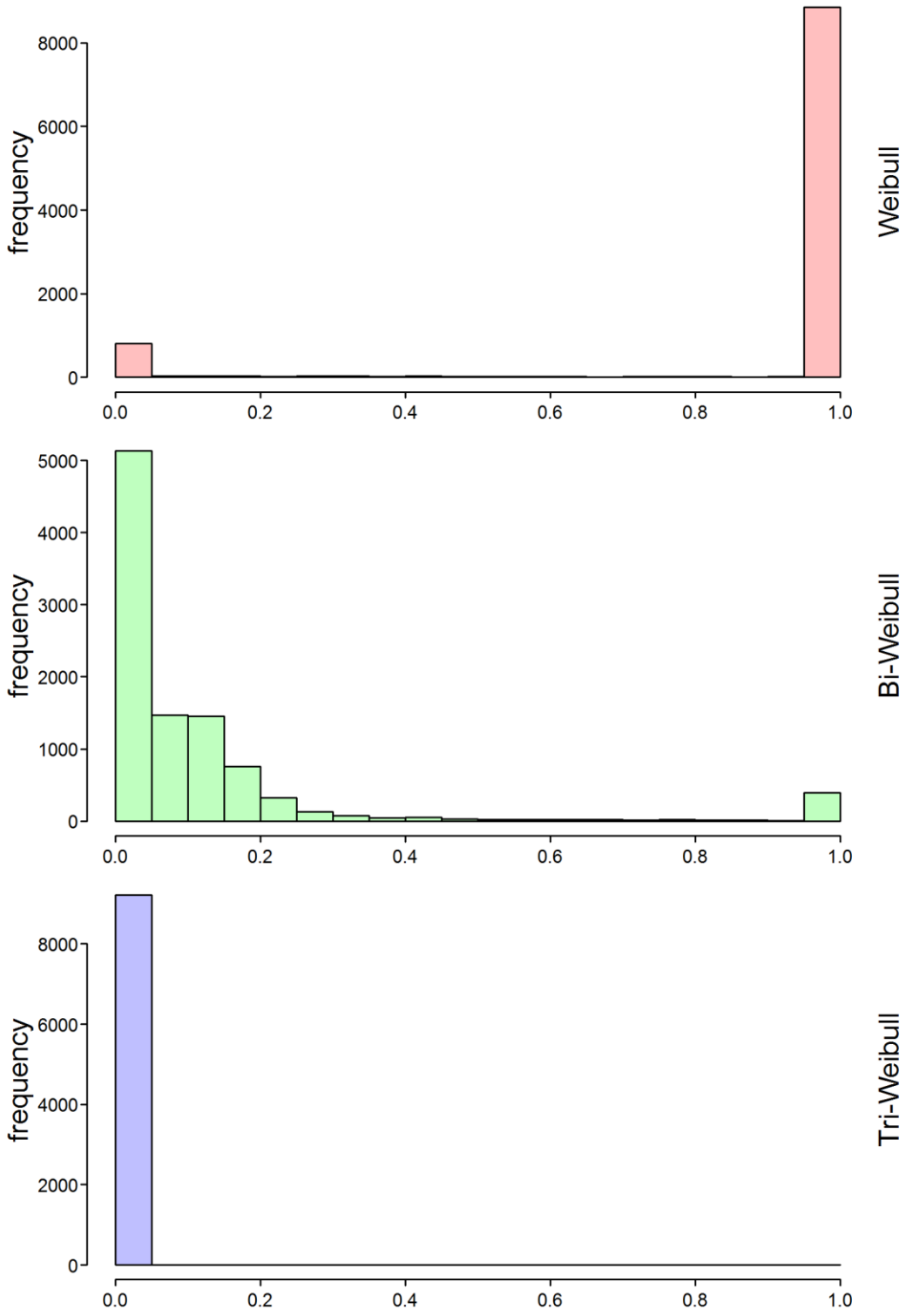

Figure 35 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 10$ systems, quality $=$ mid)

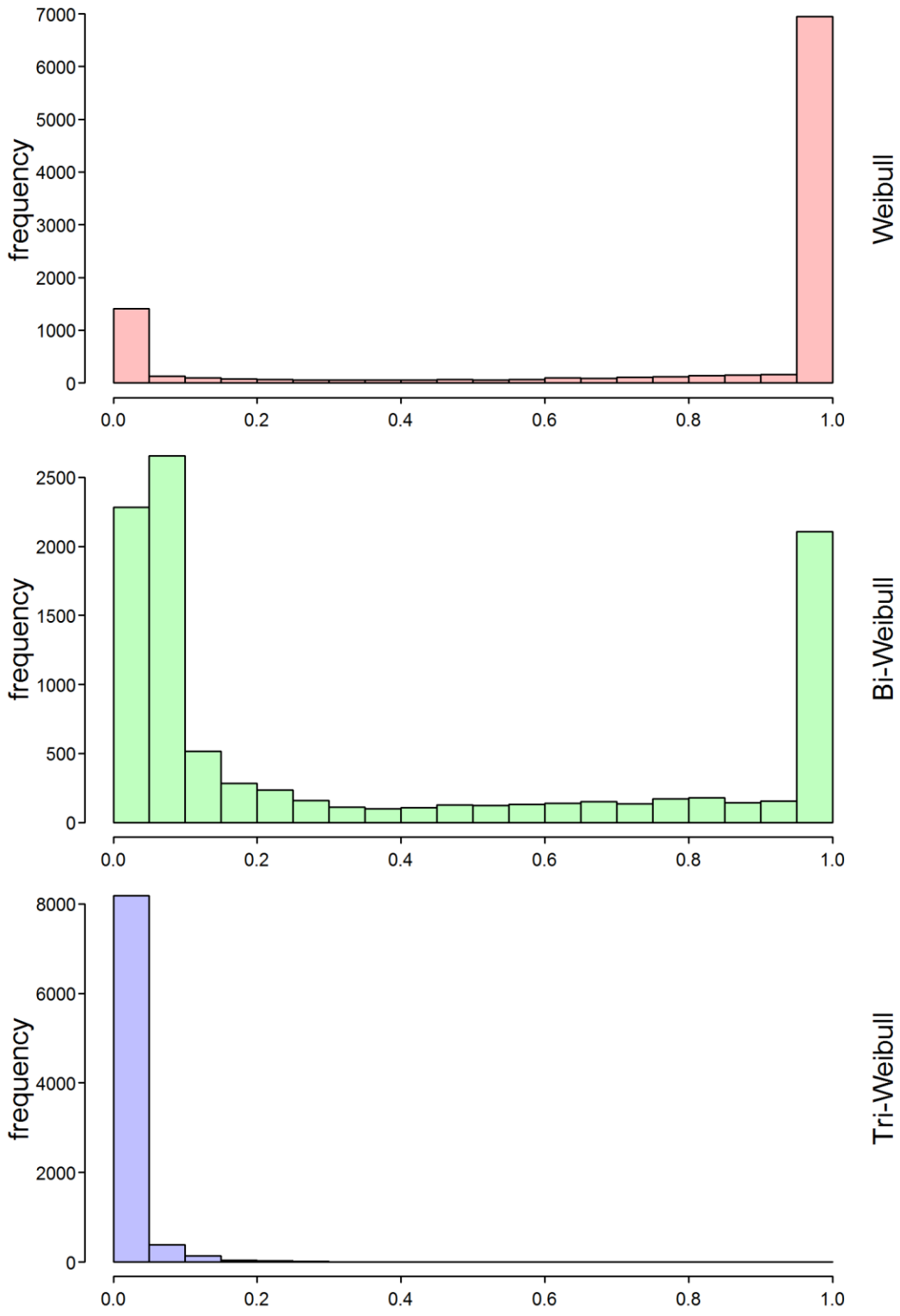

Figure 36 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 20$ systems, quality $=$ mid)

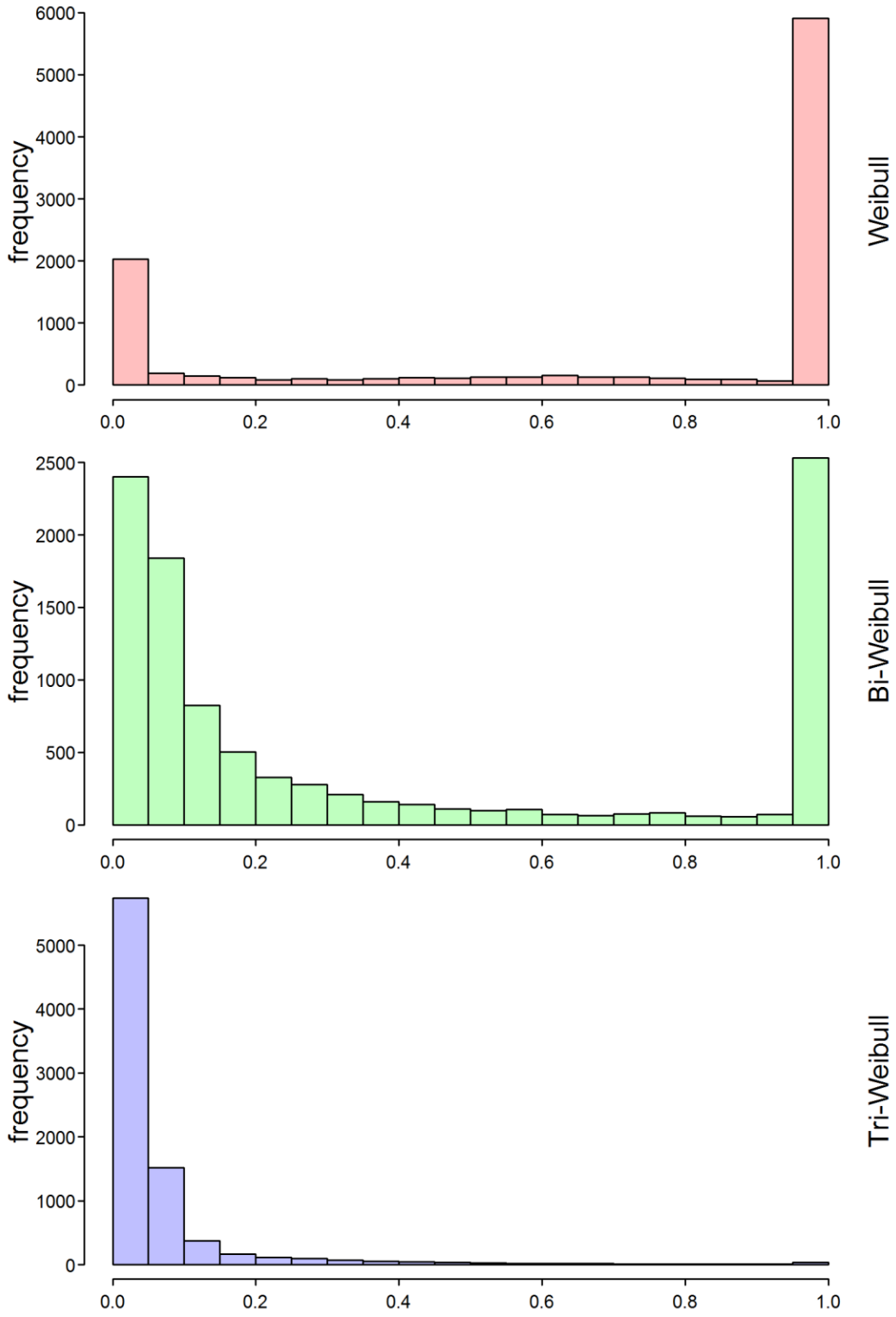

Figure 37 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 40$ systems, quality $=$ mid)

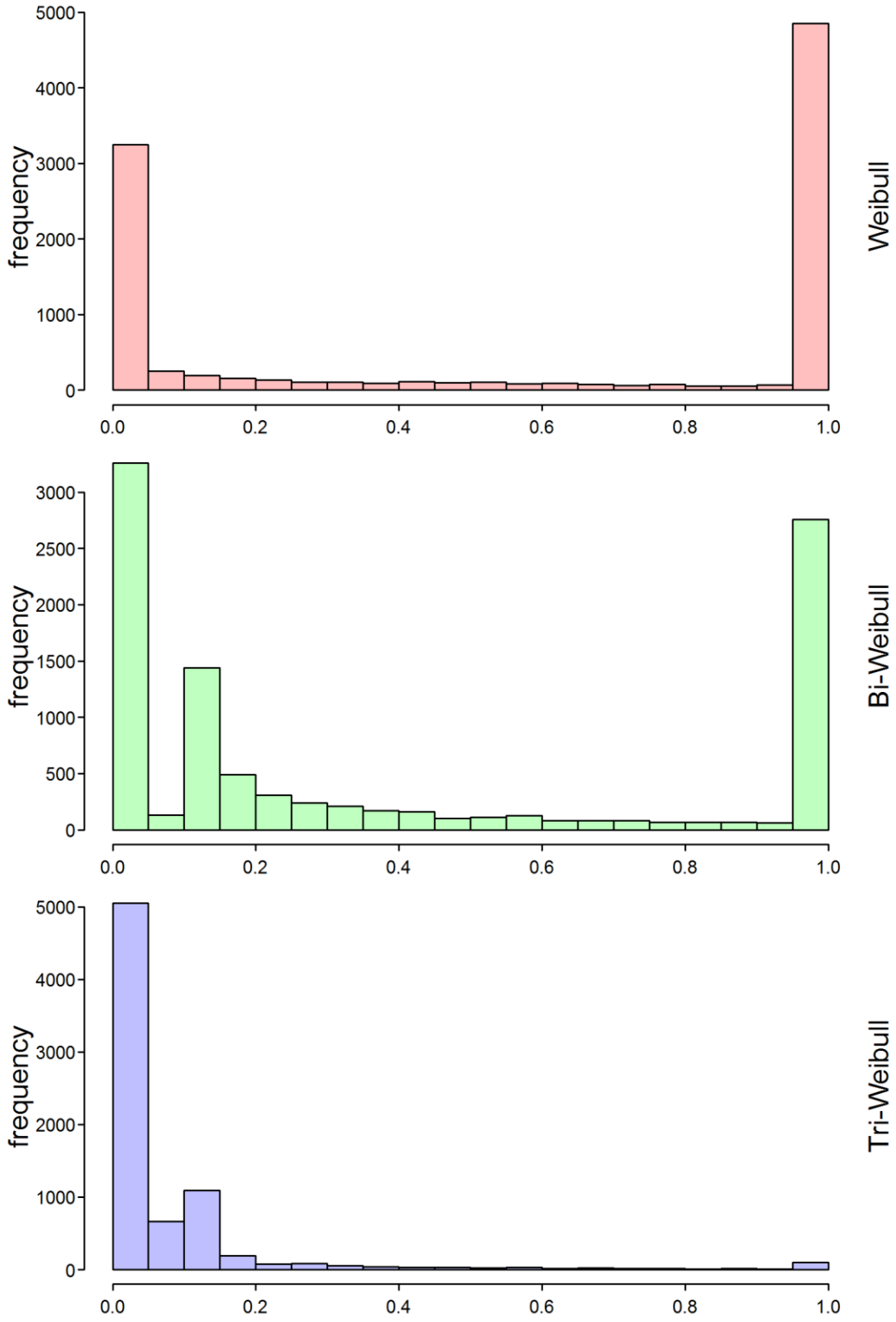

Figure 38 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 100$ systems, quality $=$ mid)

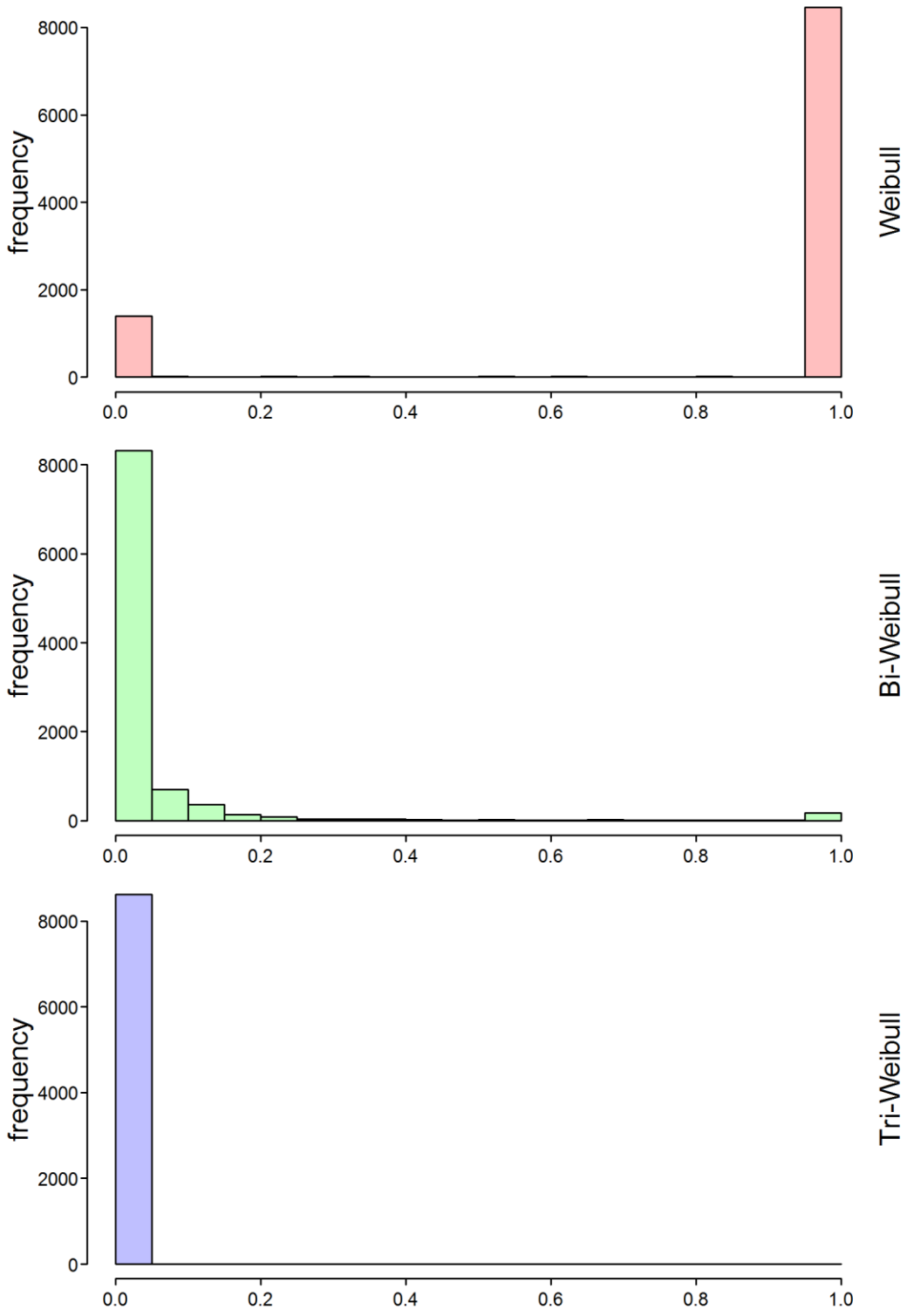

Figure 39 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 10$ systems, quality $=$ high)

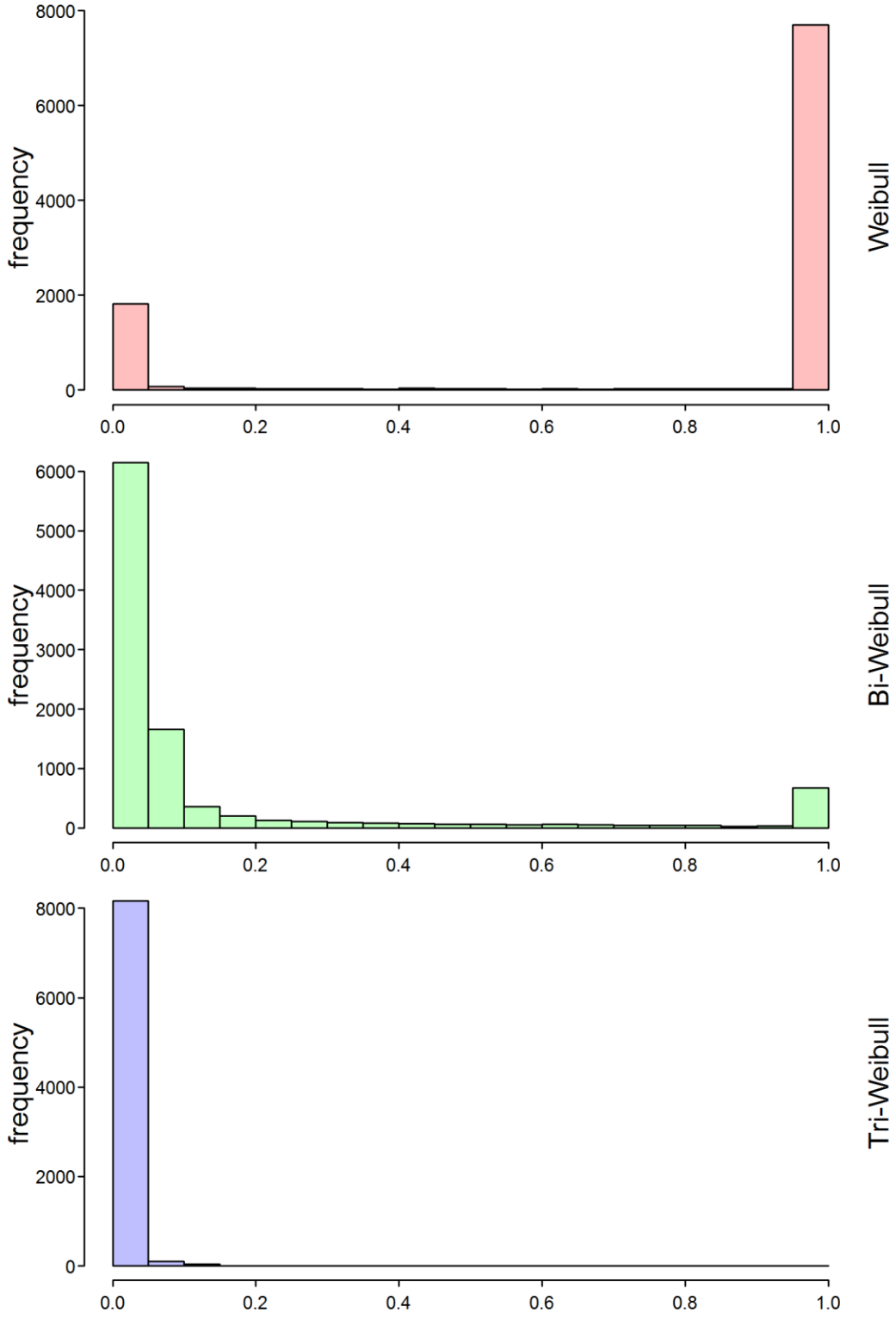

Figure 40 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 20$ systems, quality $=$ high)

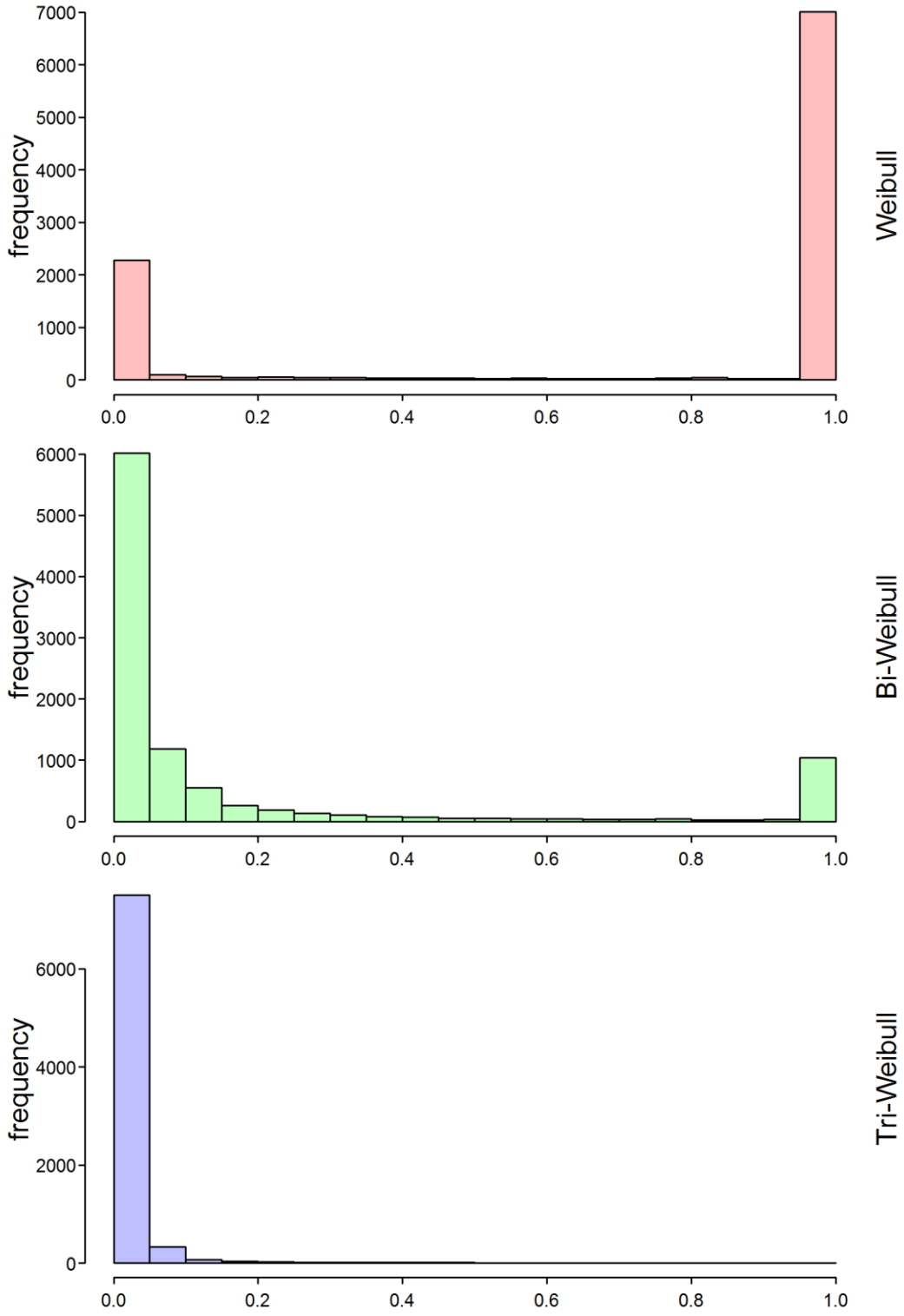

Figure 41 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 40$ systems, quality $=$ high)

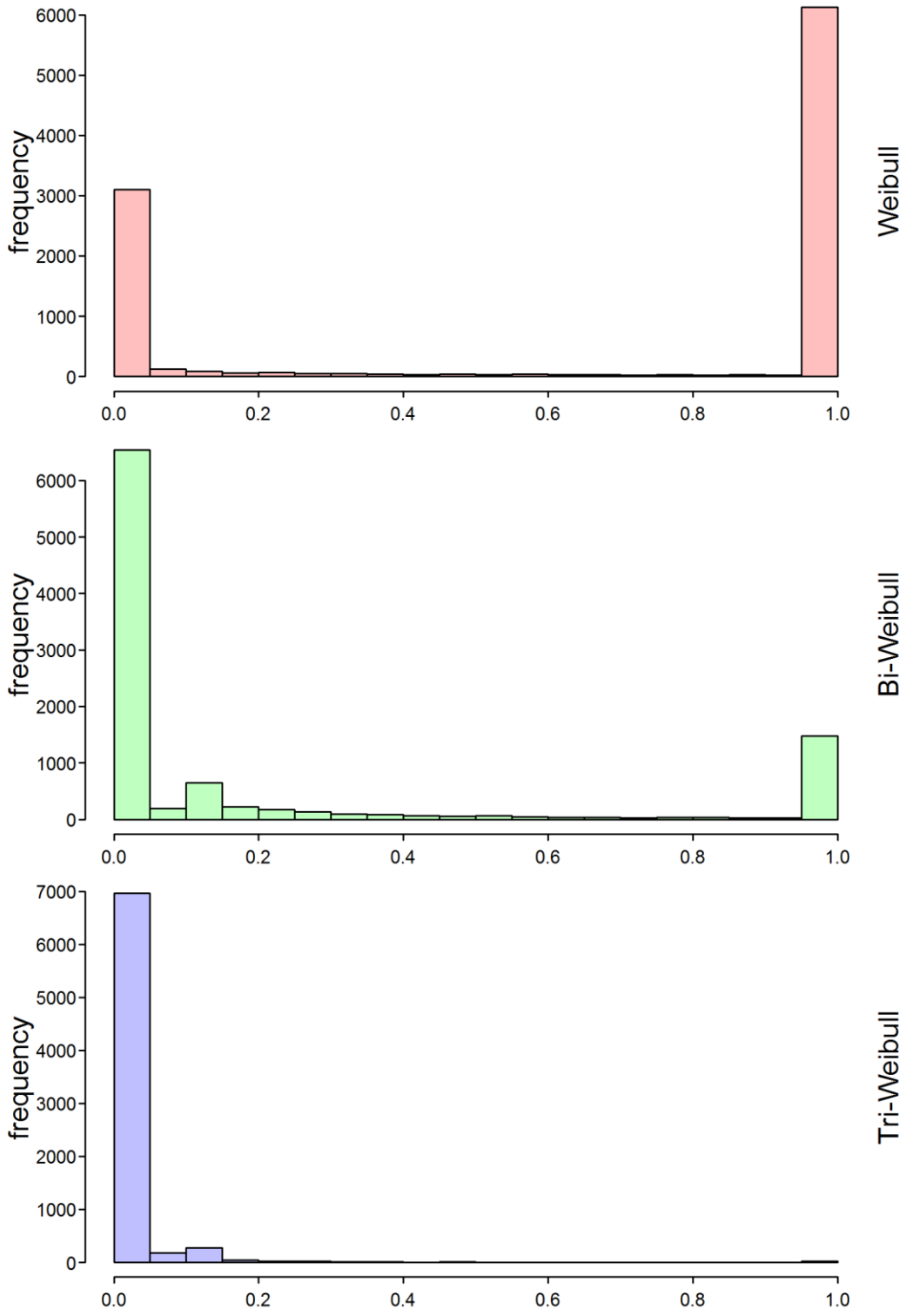

Figure 42 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 5$ modes, $n = 100$ systems, quality = high)

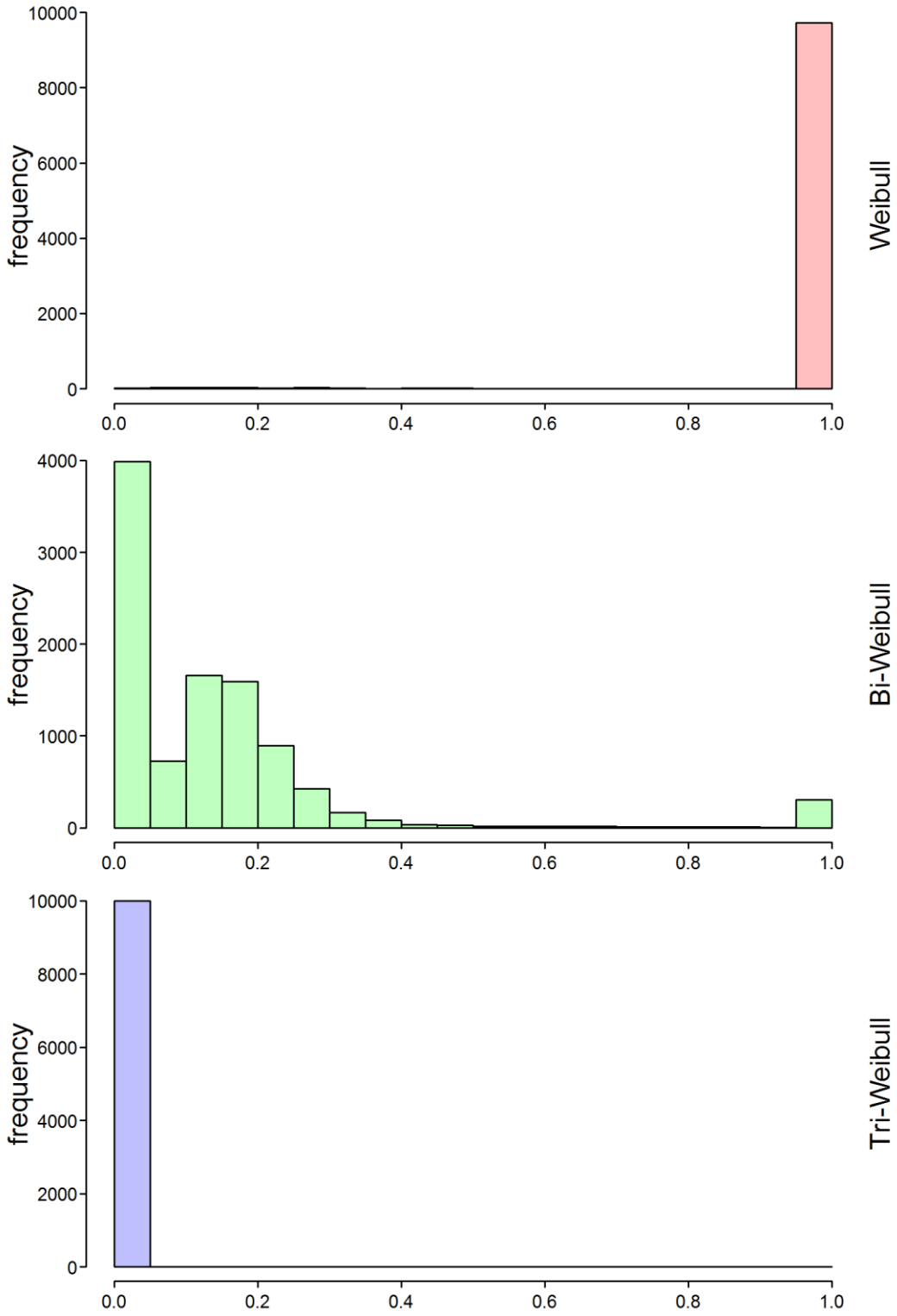

Figure 43 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 10$ systems, quality $=$ low)

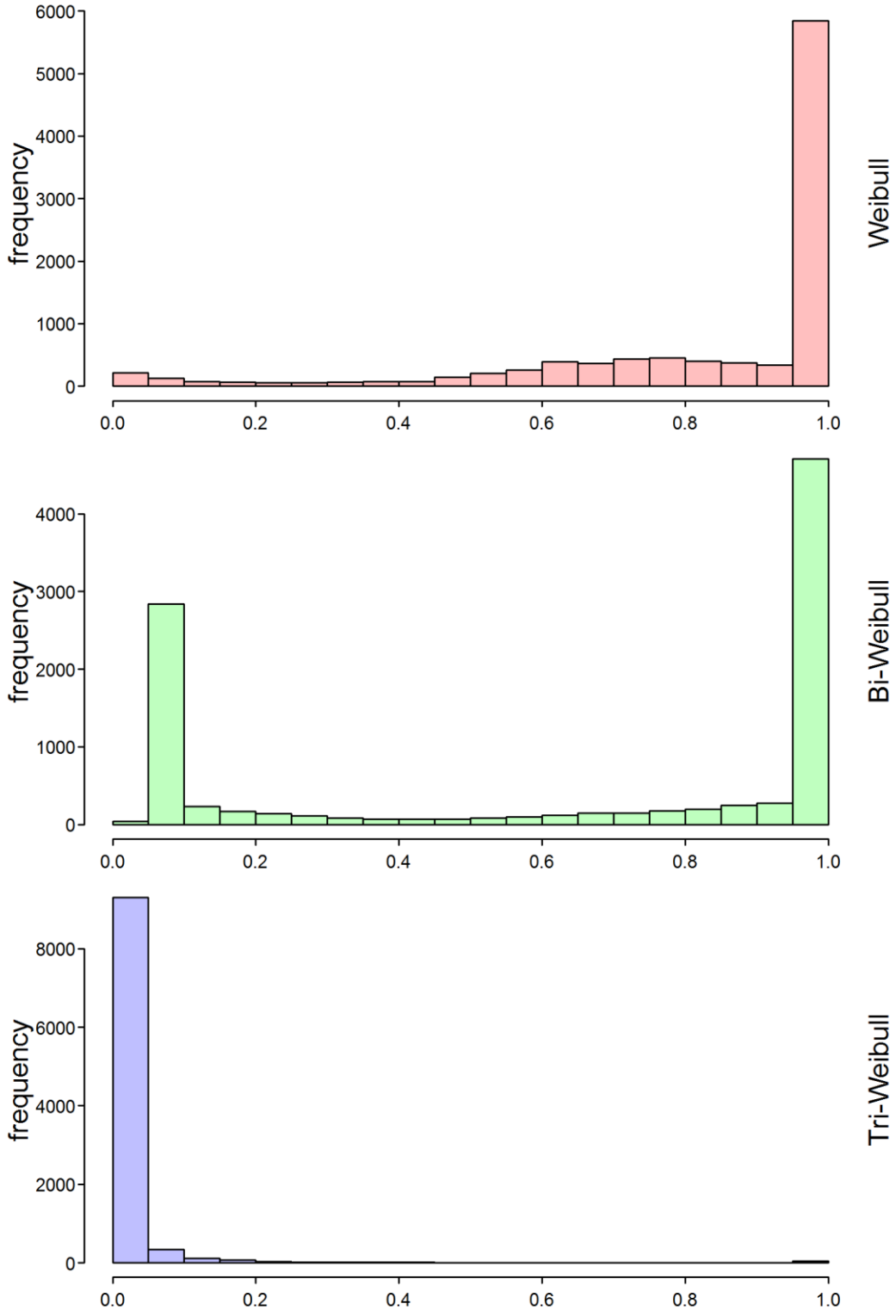

Figure 44 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 20$ systems, quality $=$ low)

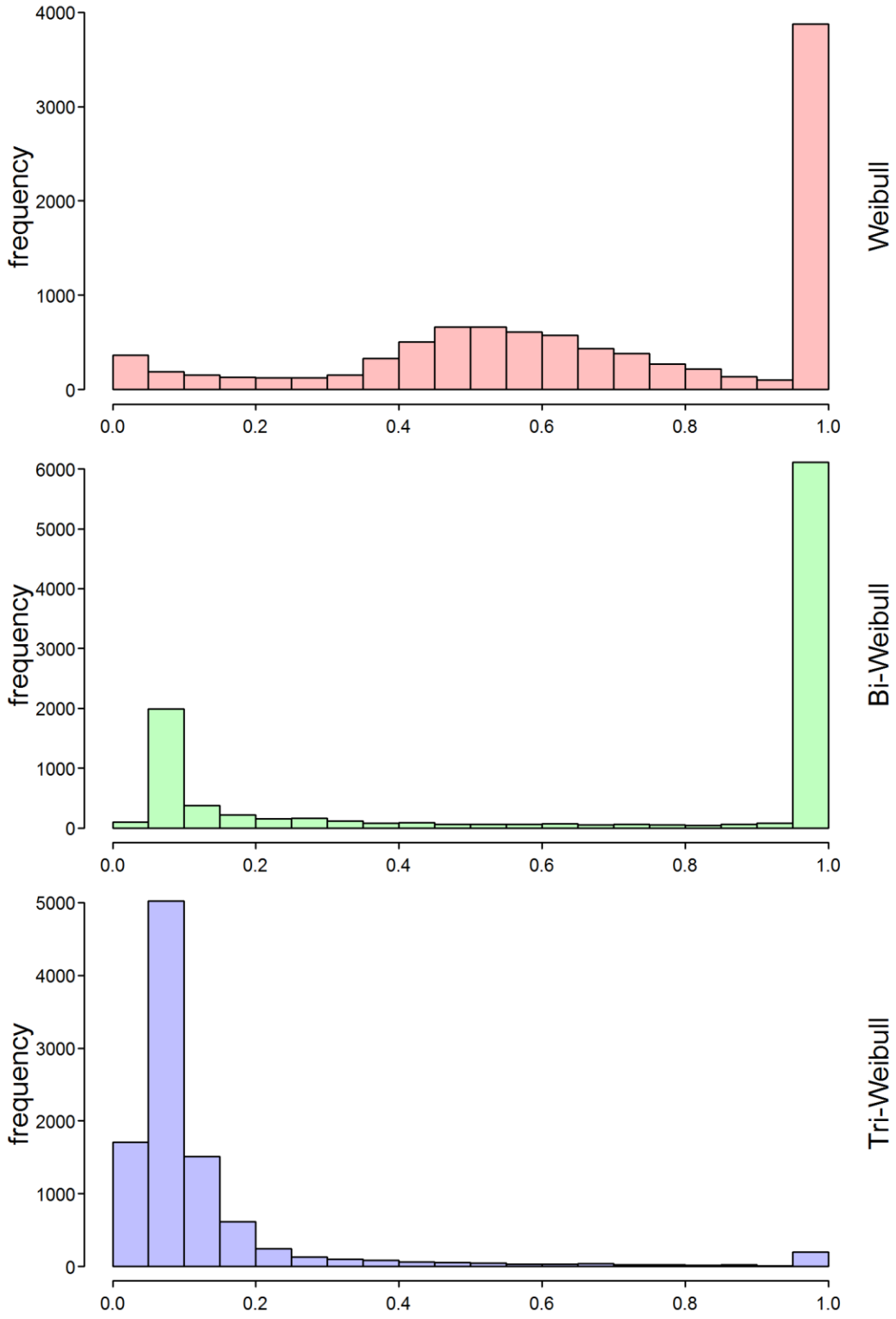

Figure 45 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 40$ systems, quality $=$ low)

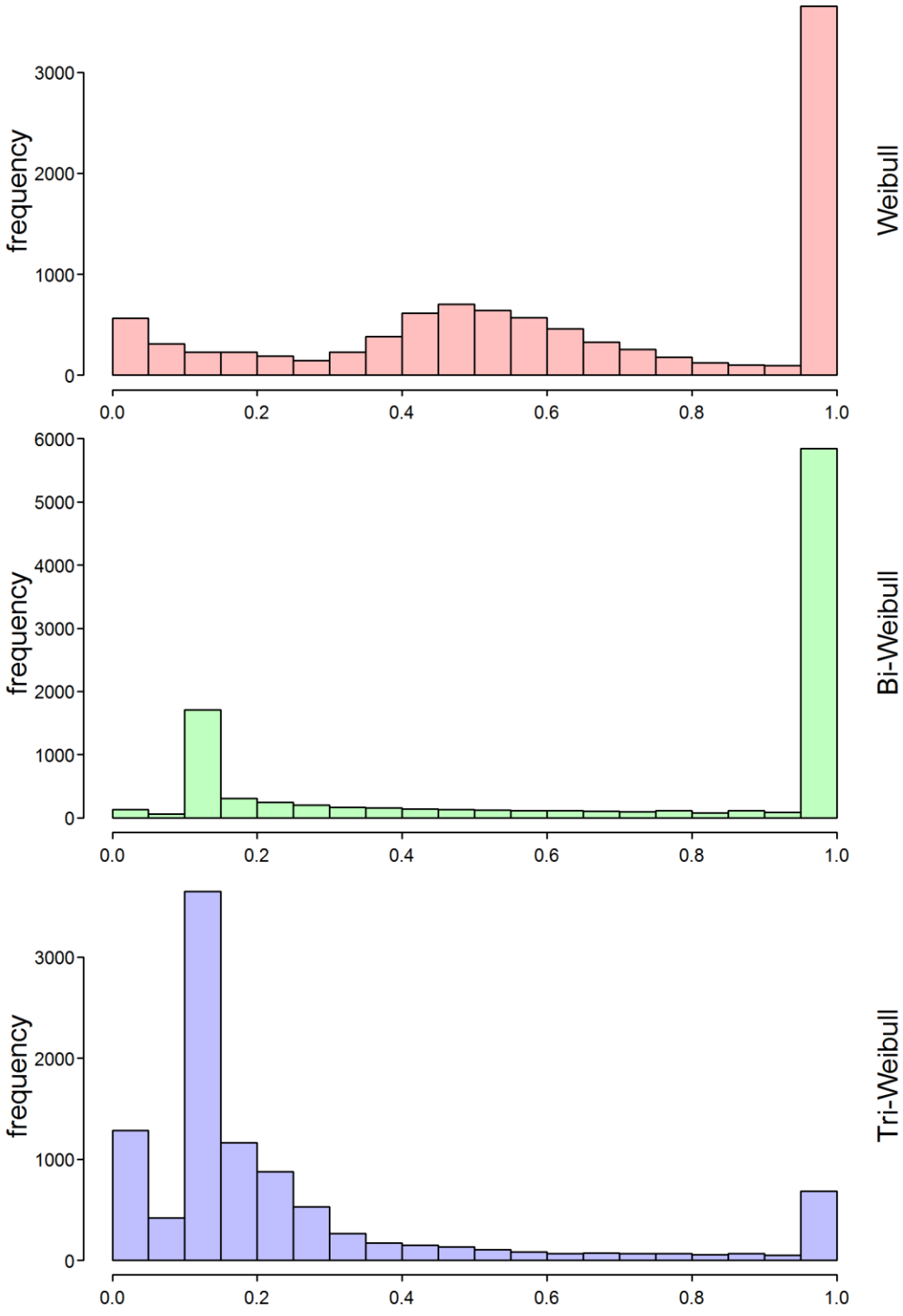

Figure 46 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 100$ systems, quality $=$ low)

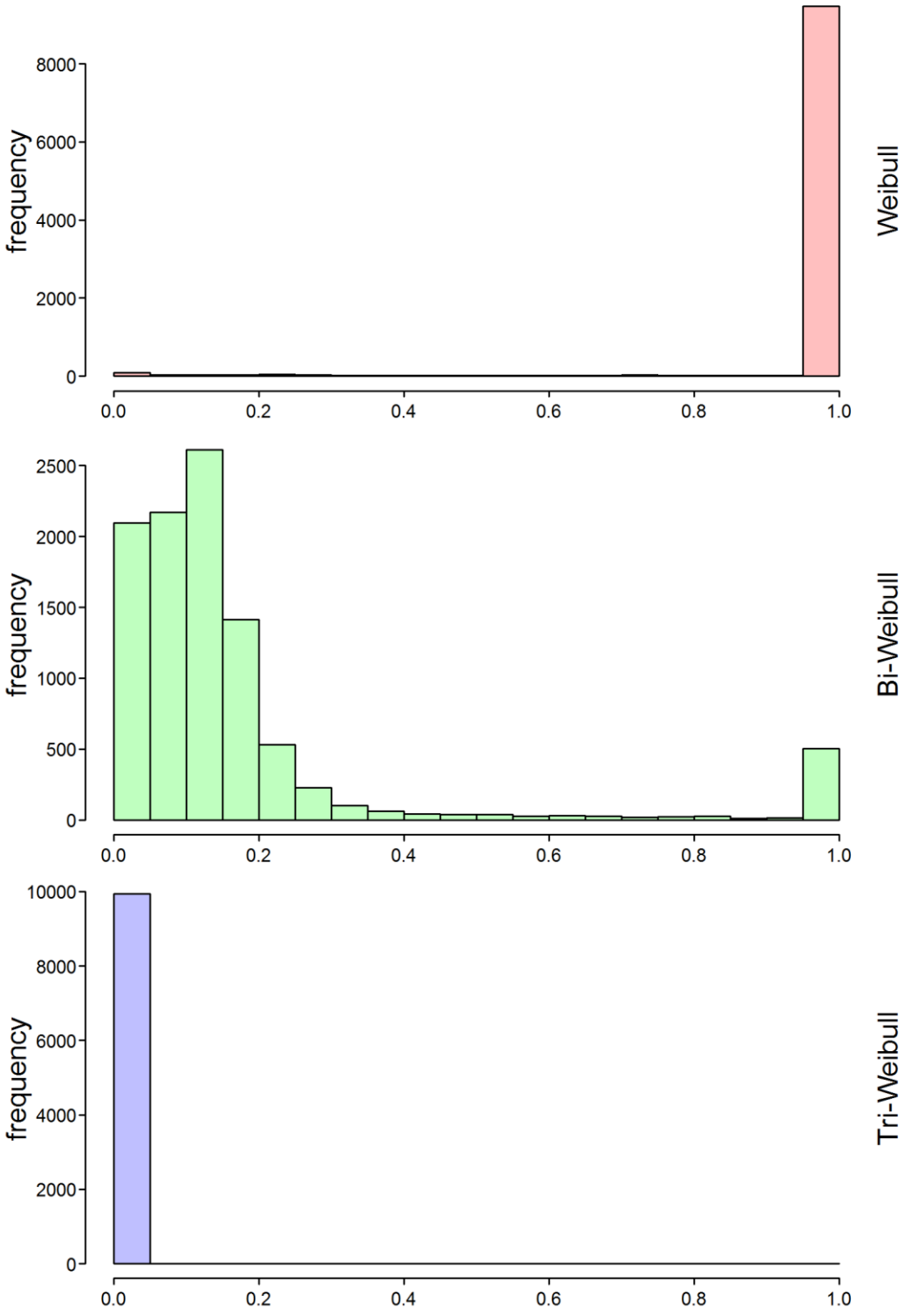

Figure 47 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 10$ systems, quality $=$ mid)

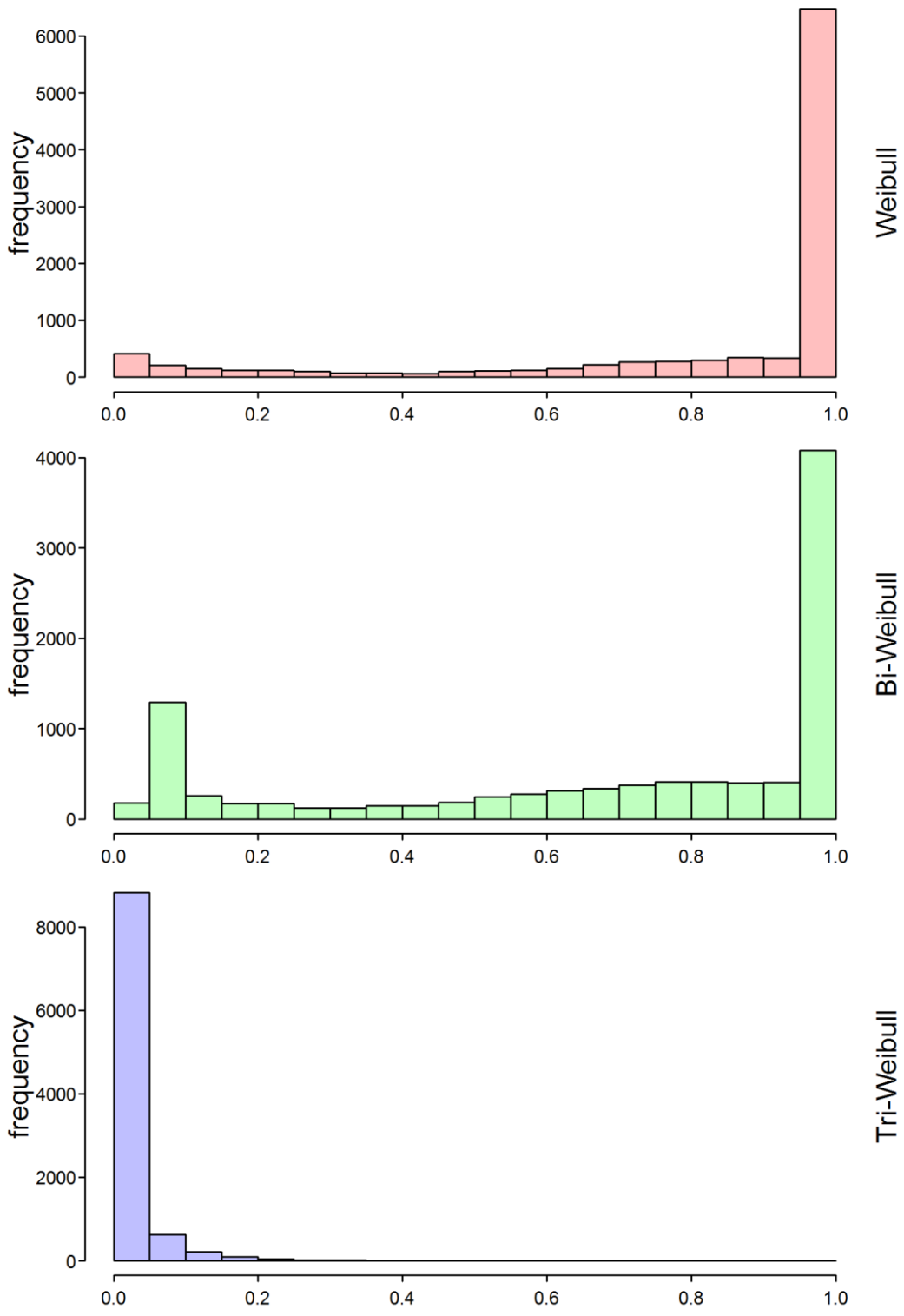

Figure 48 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 20$ systems, quality $=$ mid)

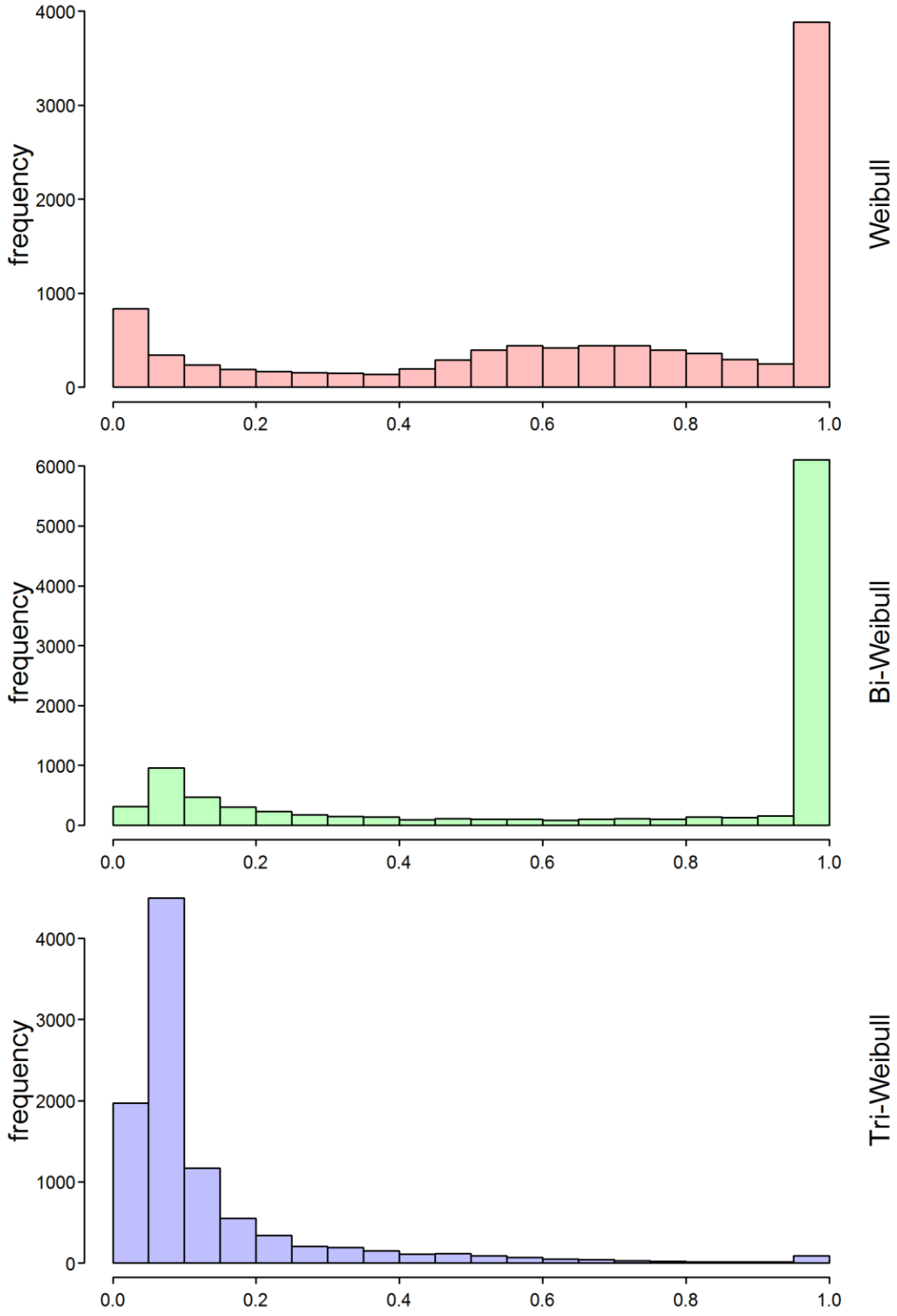

Figure 49 – Histogram of AICc values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 40$ systems, quality $=$ mid)

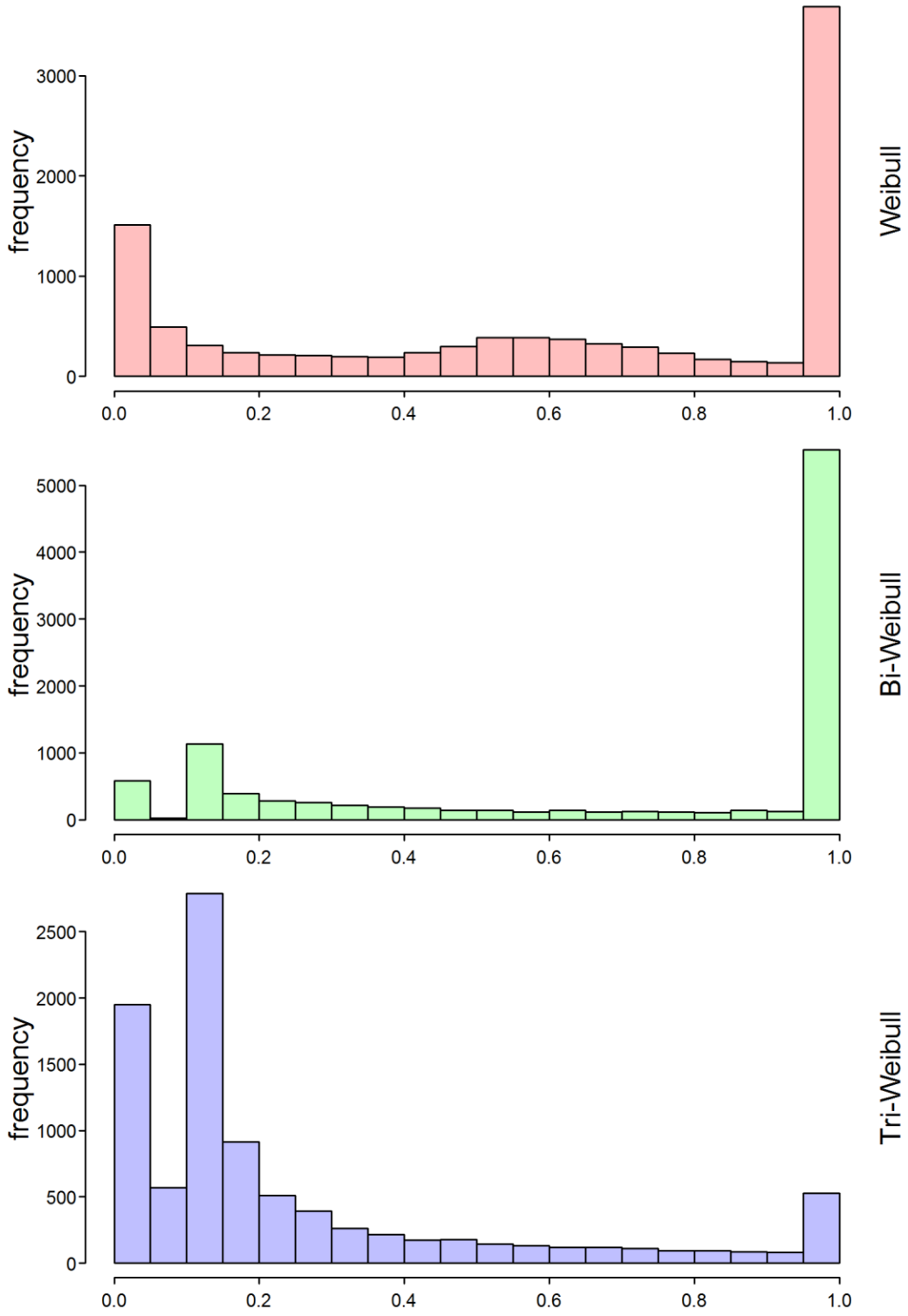

Figure 50 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 100$ systems, quality $=$ mid)

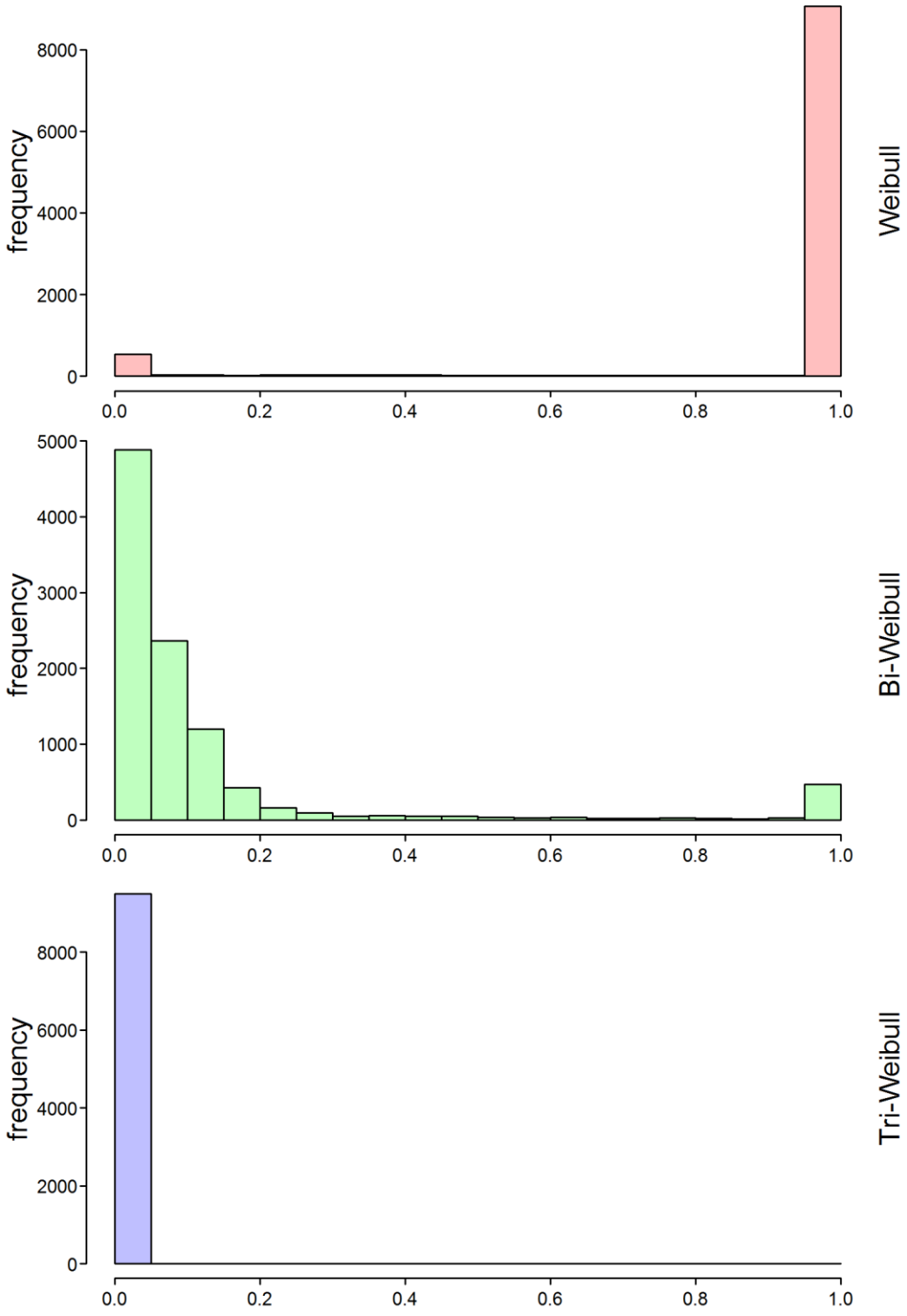

Figure 51 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 10$ systems, quality $=$ high)

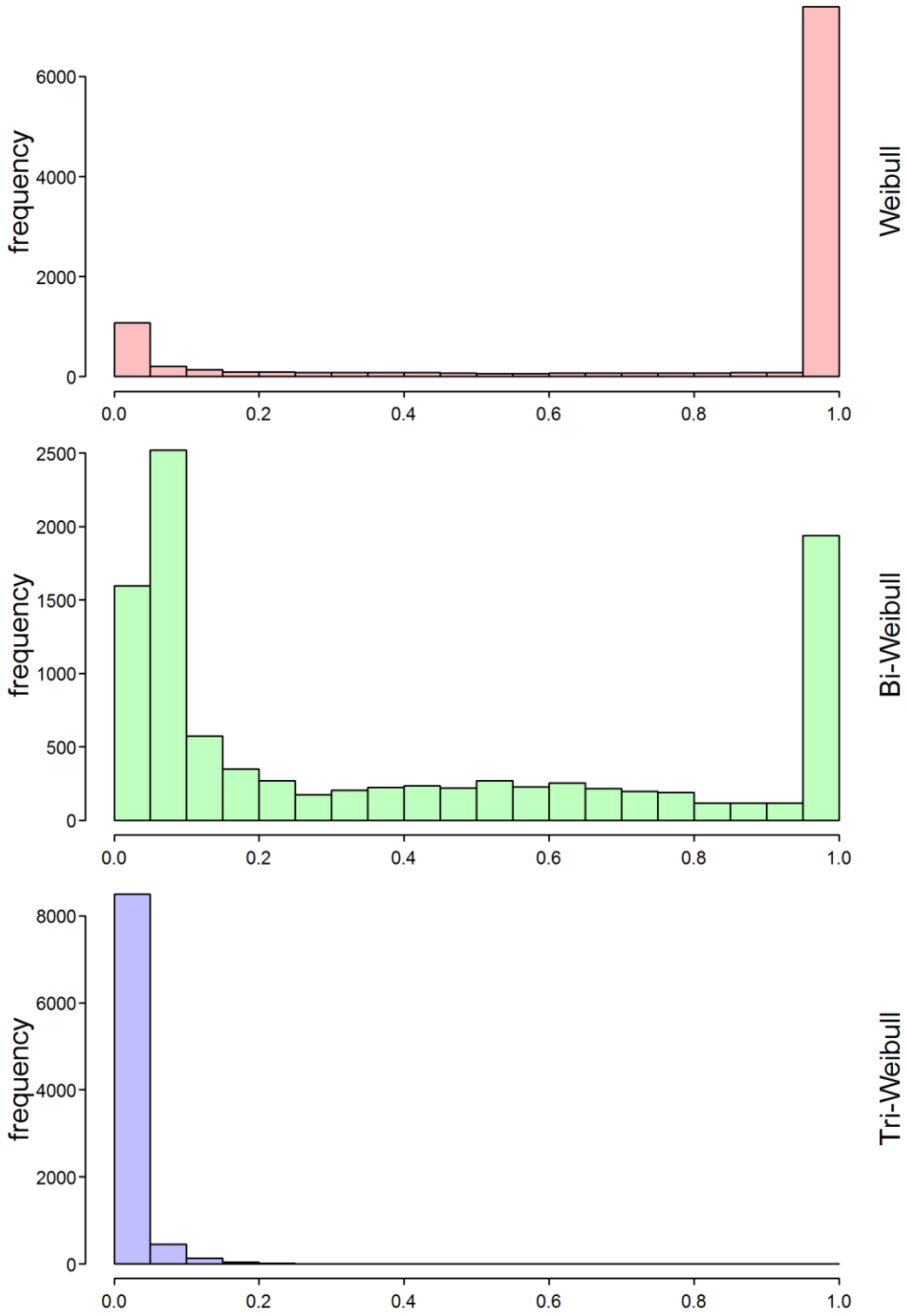

Figure 52 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 20$ systems, quality = high)

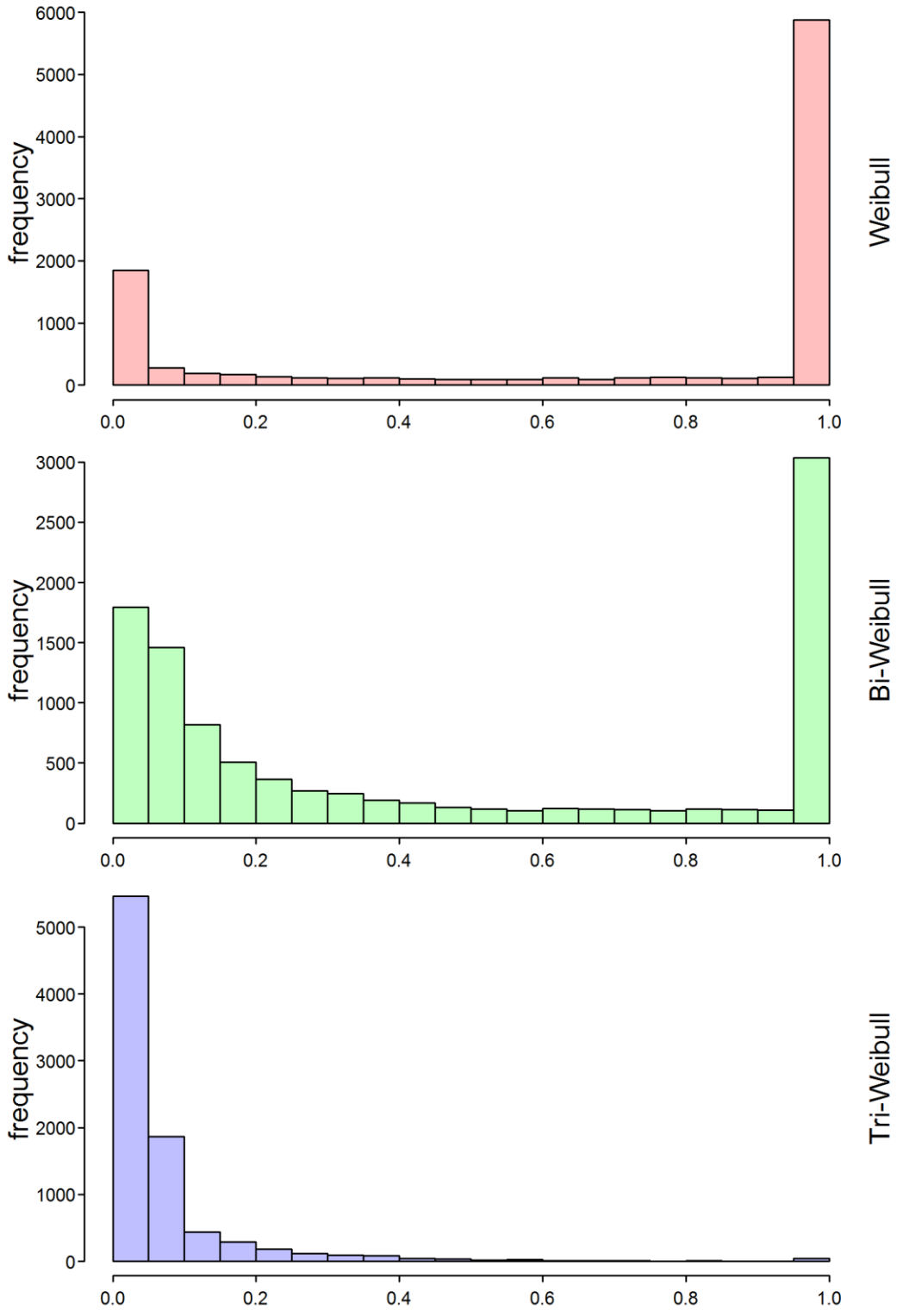

Figure 53 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 40$ systems, quality $=$ high)

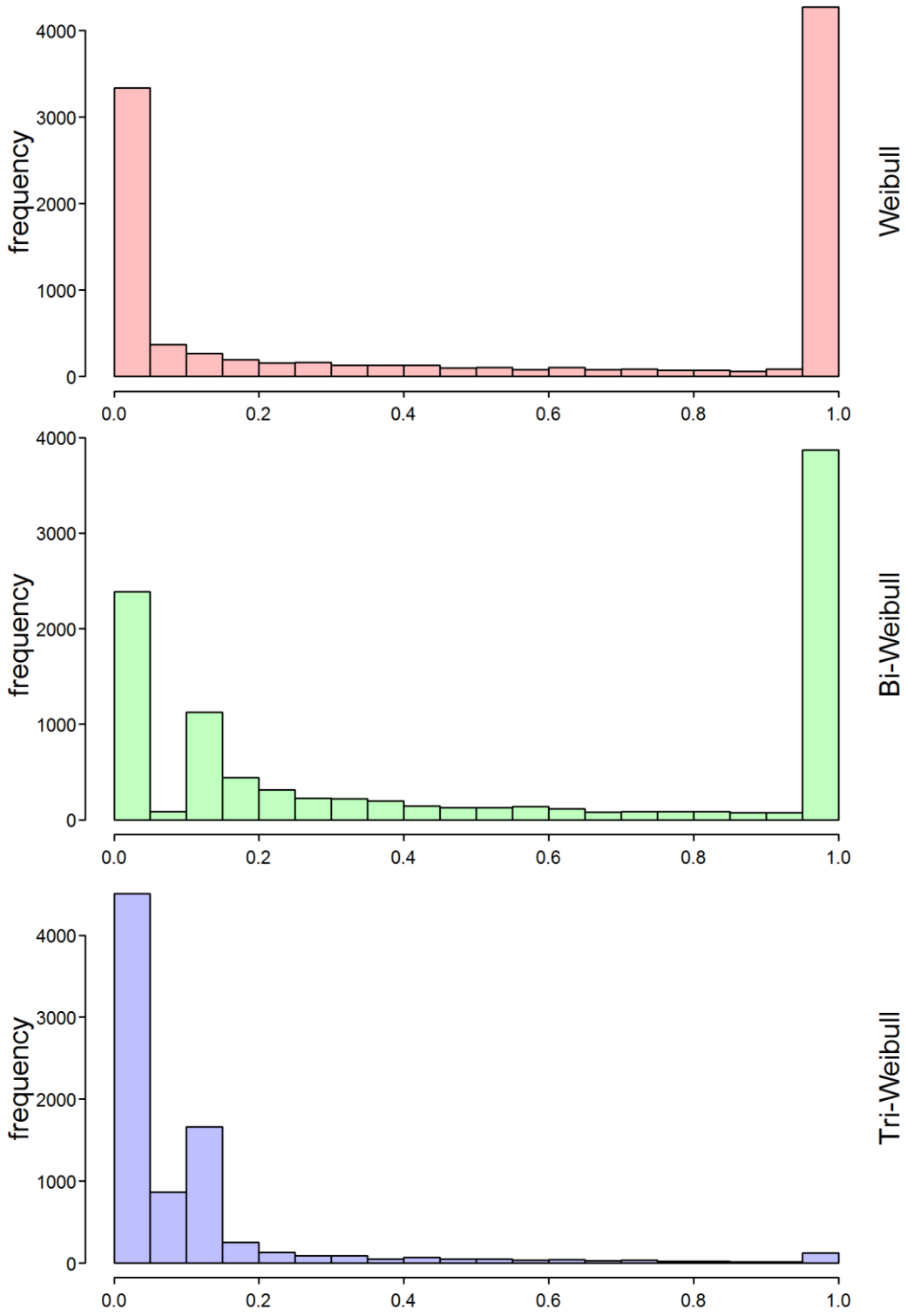

Figure 54 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 10$ modes, $n = 100$ systems, quality $=$ high)

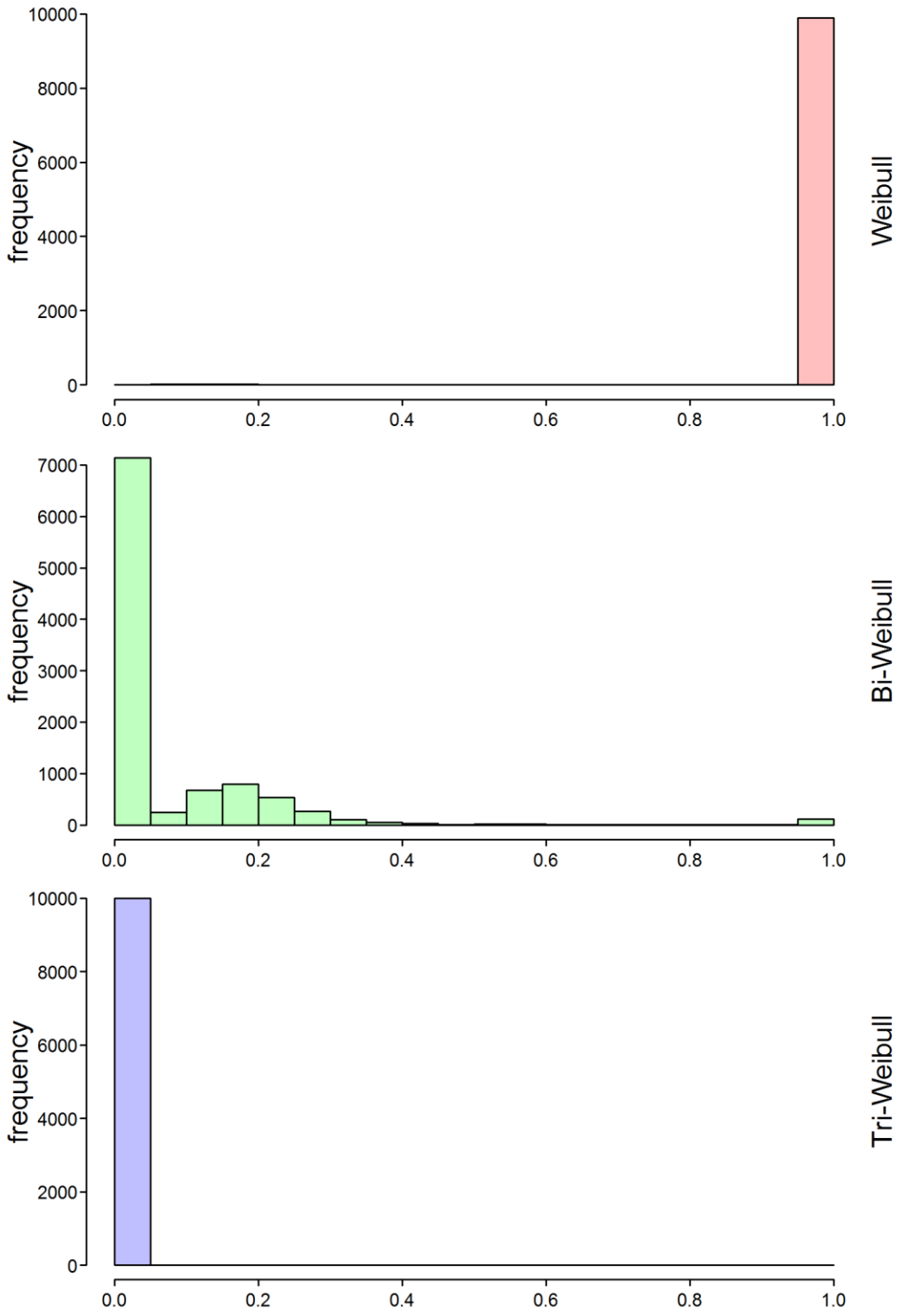

Figure 55 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 10$ systems, quality $=$ low)

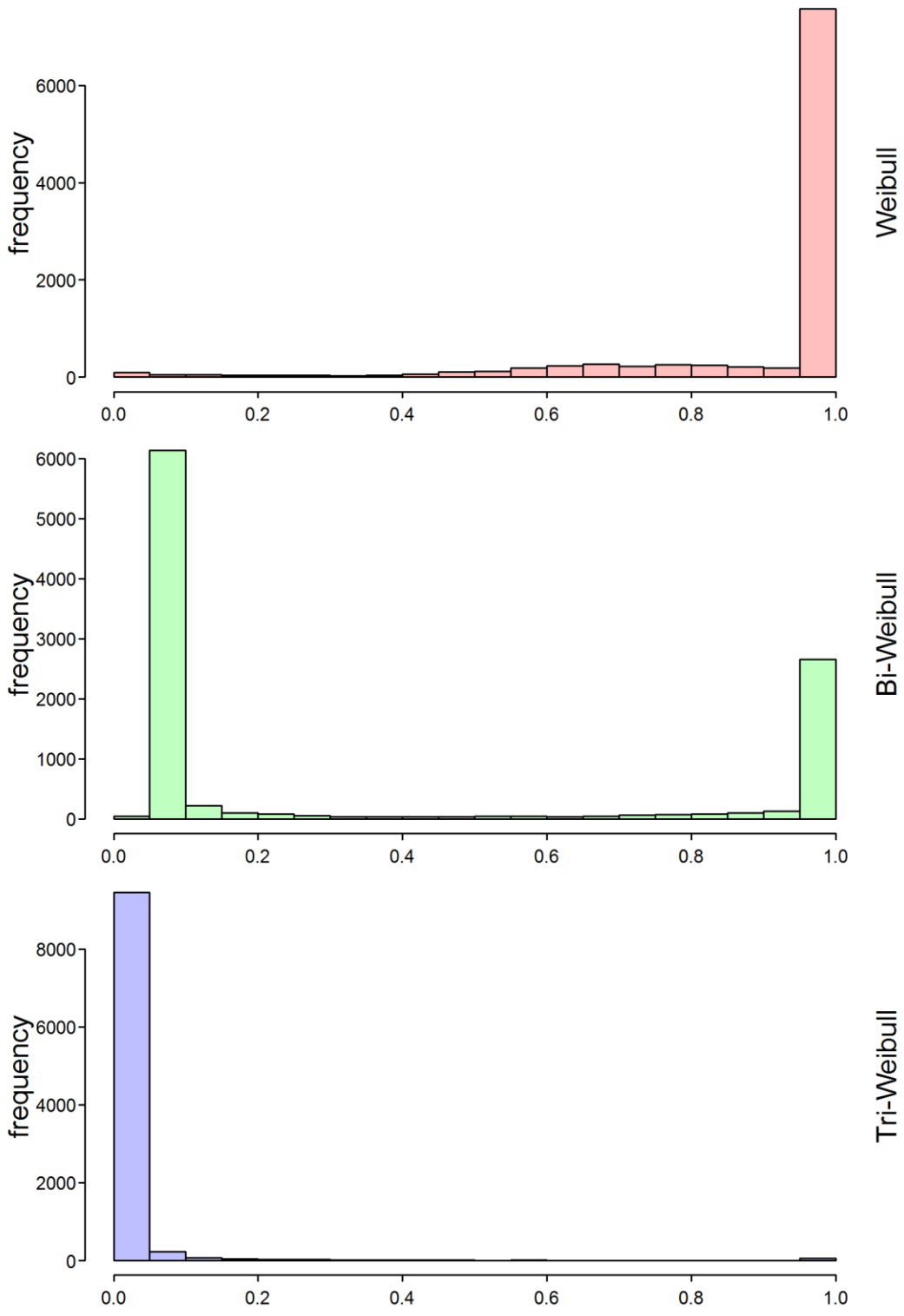

Figure 56 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 20$ systems, quality $=$ low)

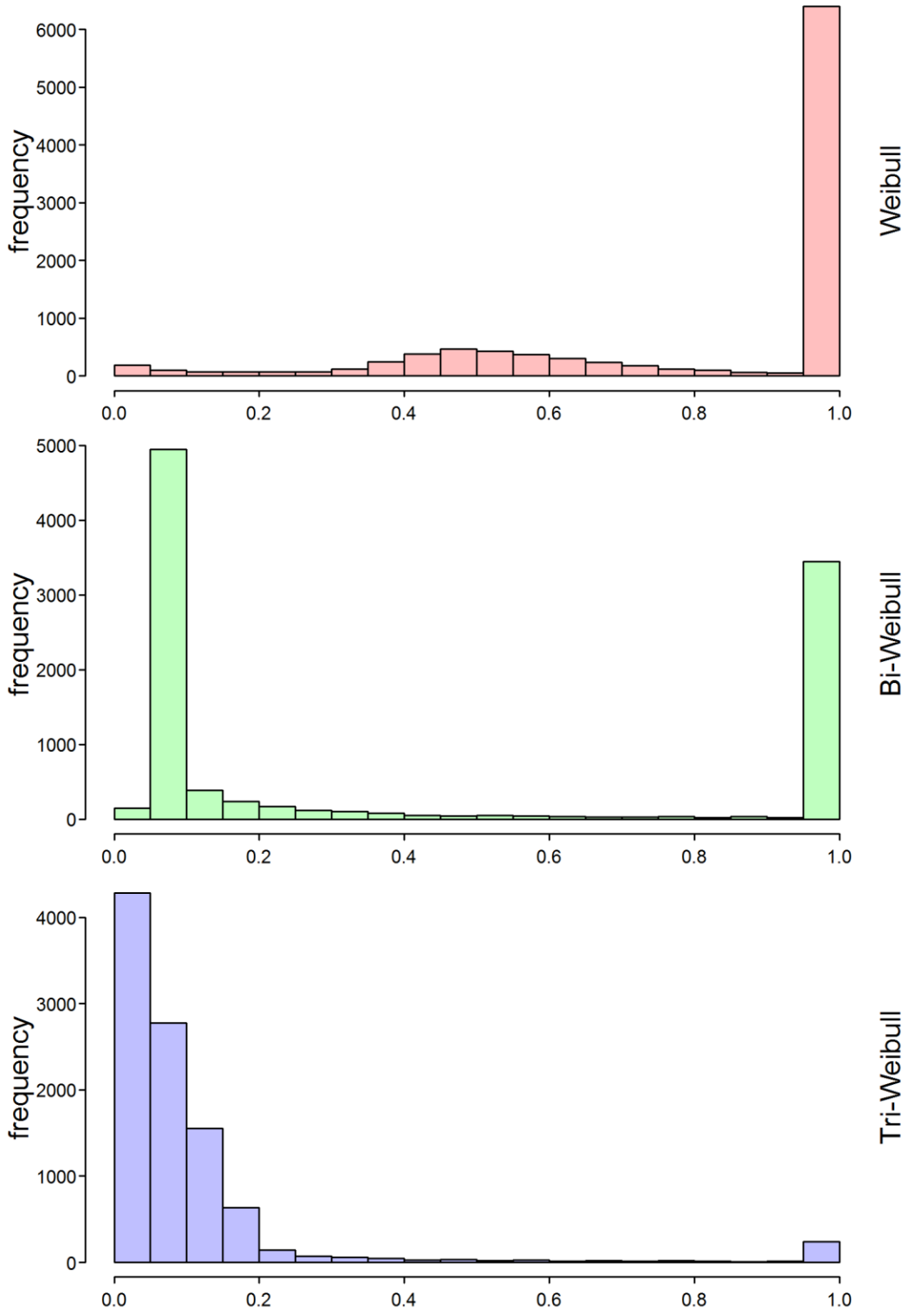

Figure 57 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 40$ systems, quality $=$ low)

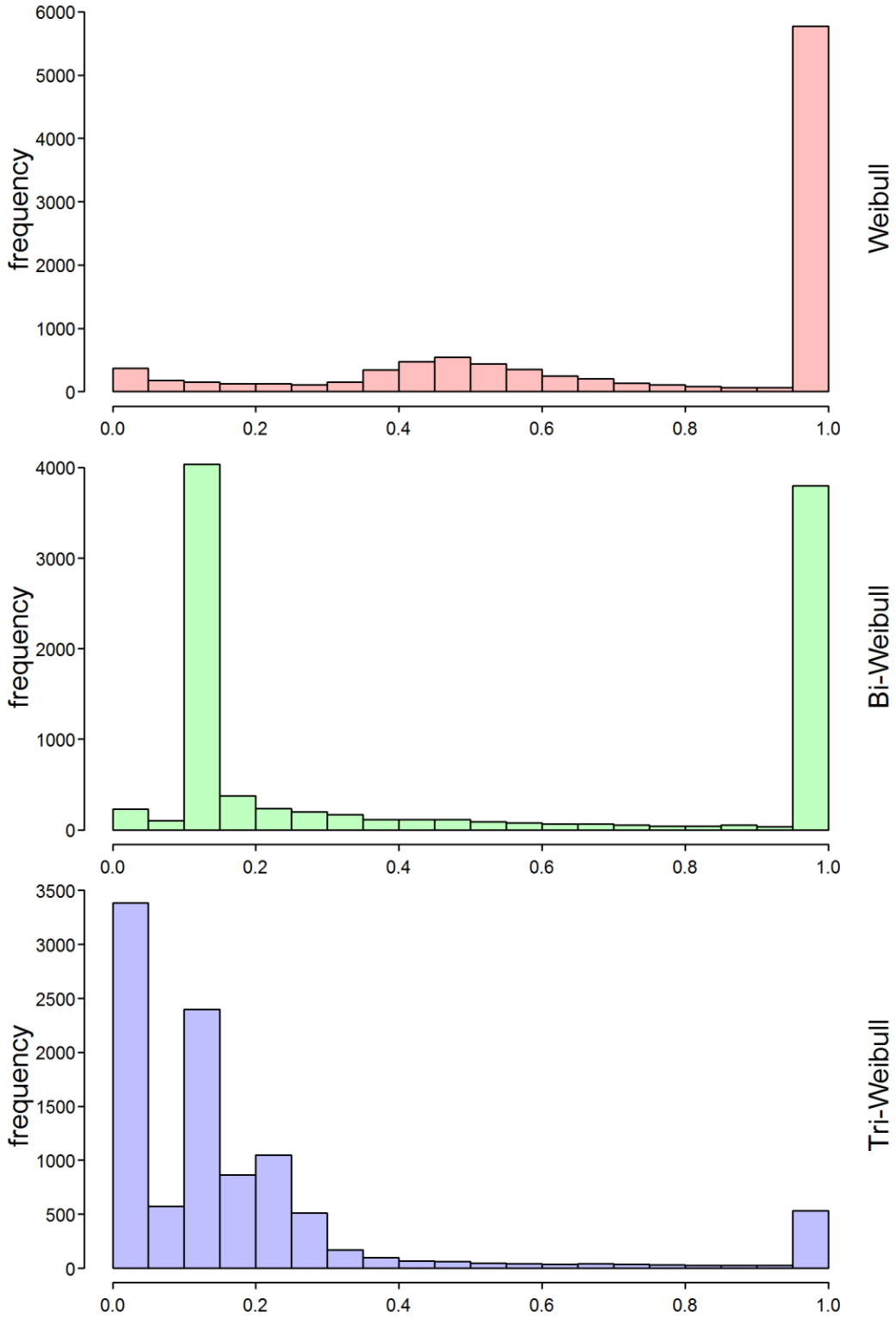

Figure 58 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 100$ systems, quality $=$ low)

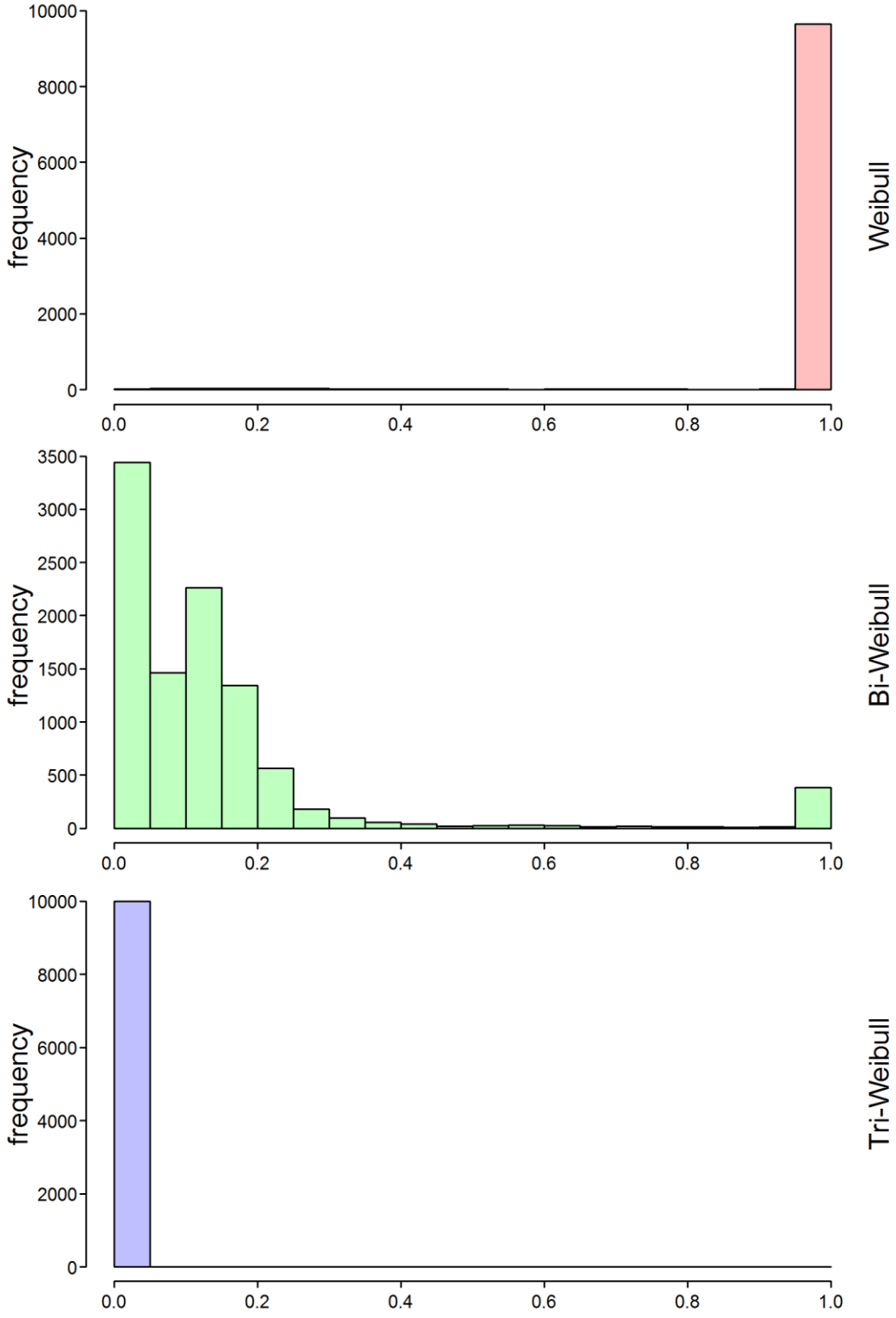

Figure 59 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 10$ systems, quality $=$ mid)

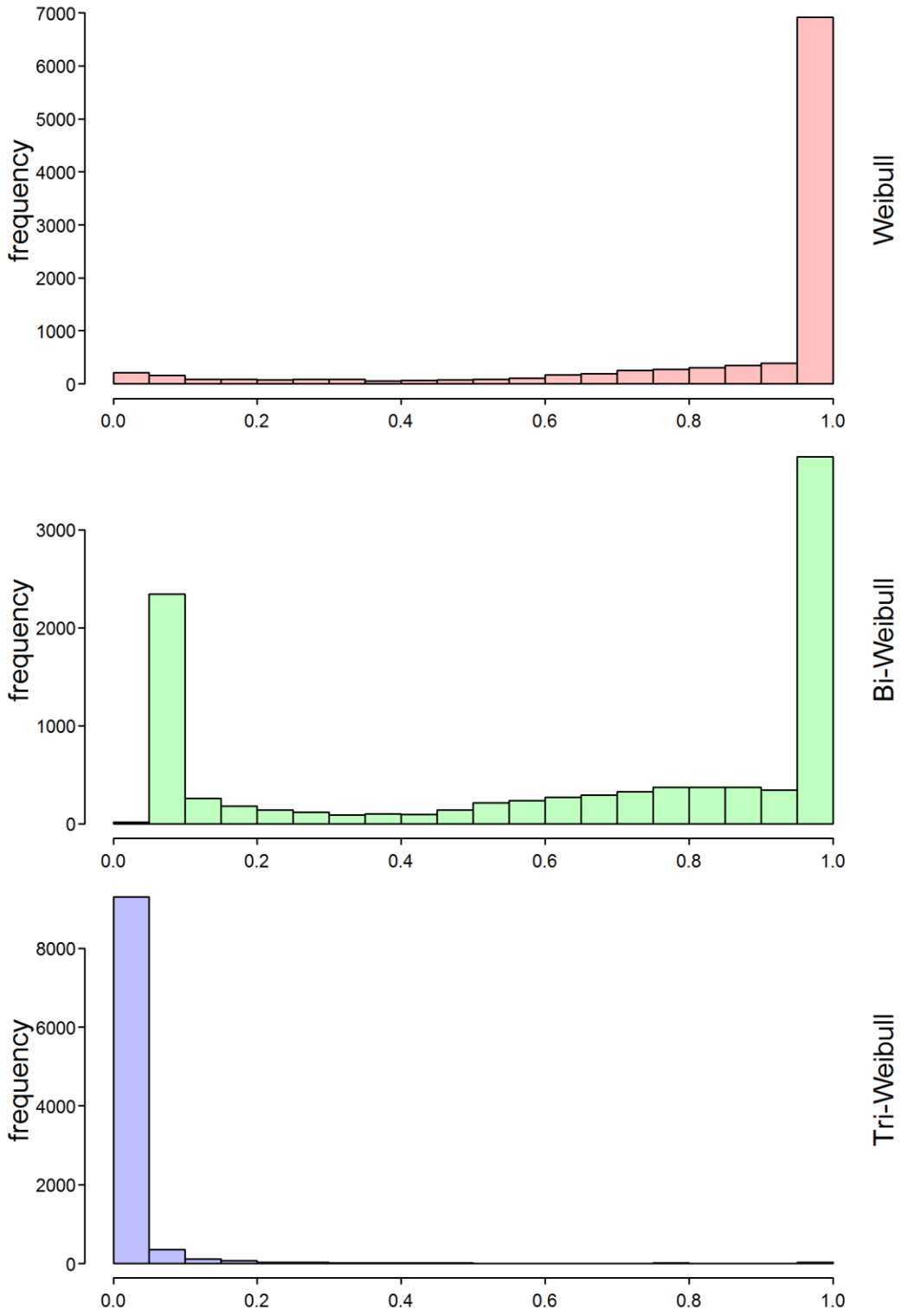

Figure 60 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 20$ systems, quality $=$ mid)

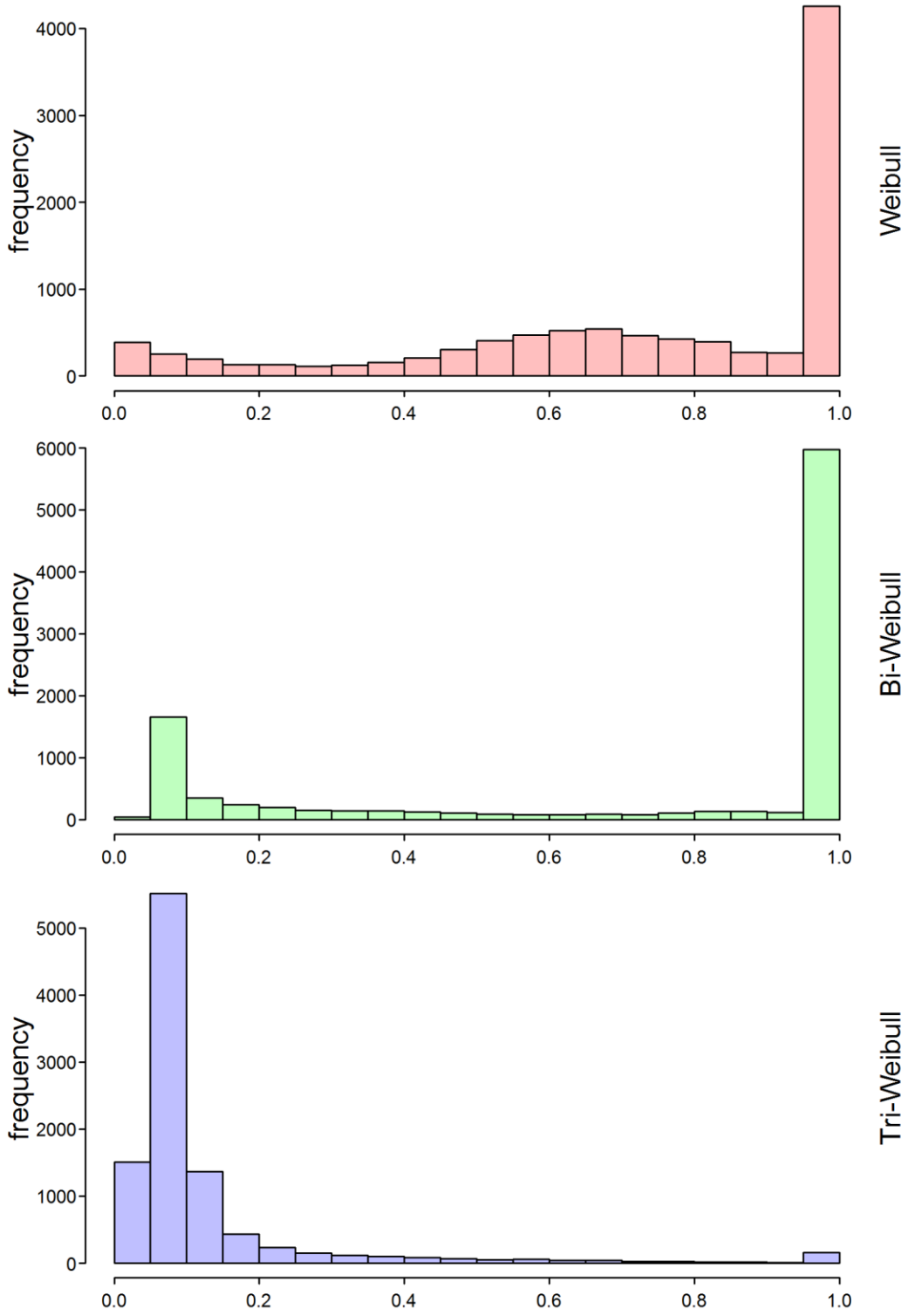

Figure 61 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 40$ systems, quality = mid)

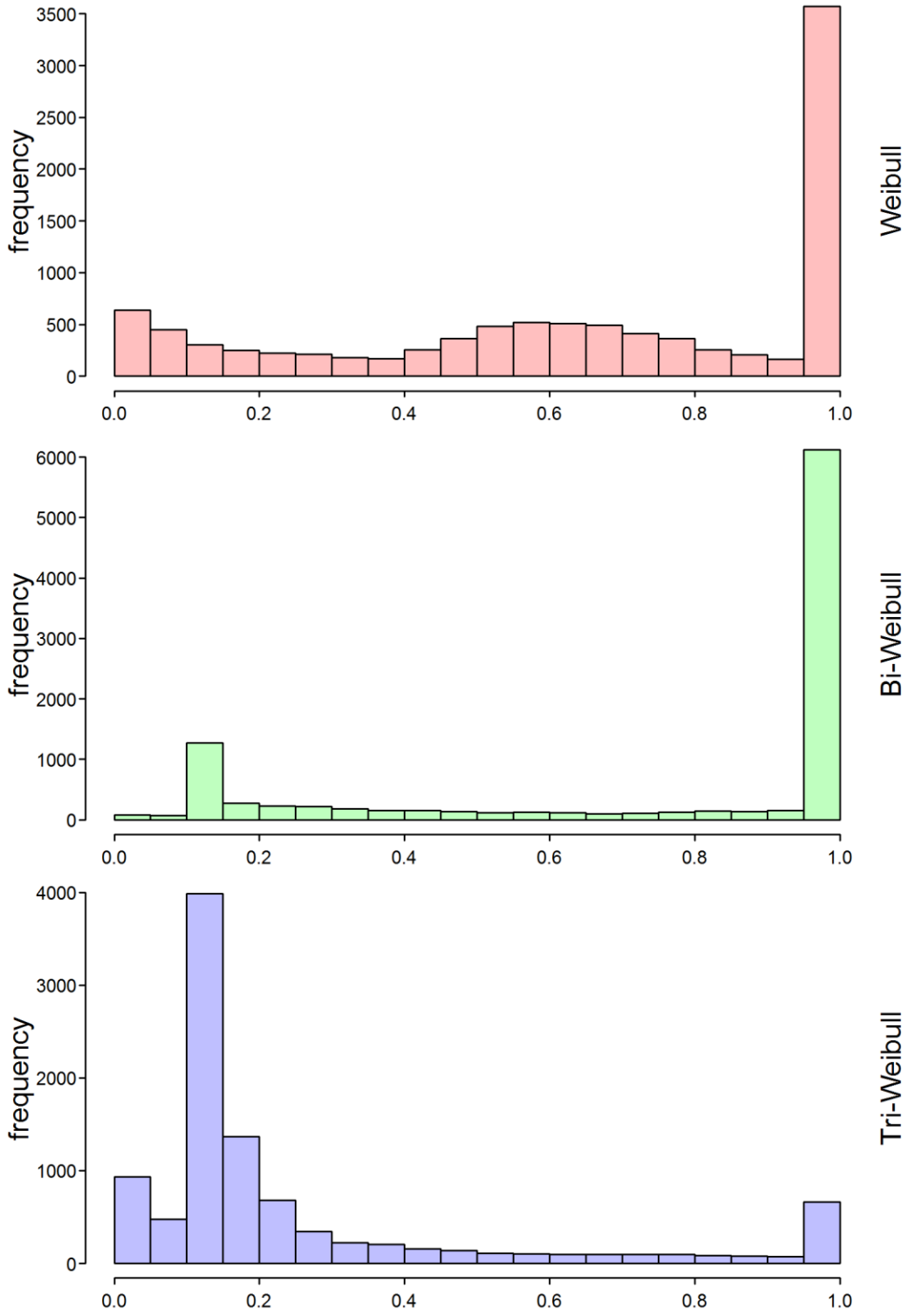

Figure 62 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 100$ systems, quality $=$ mid)

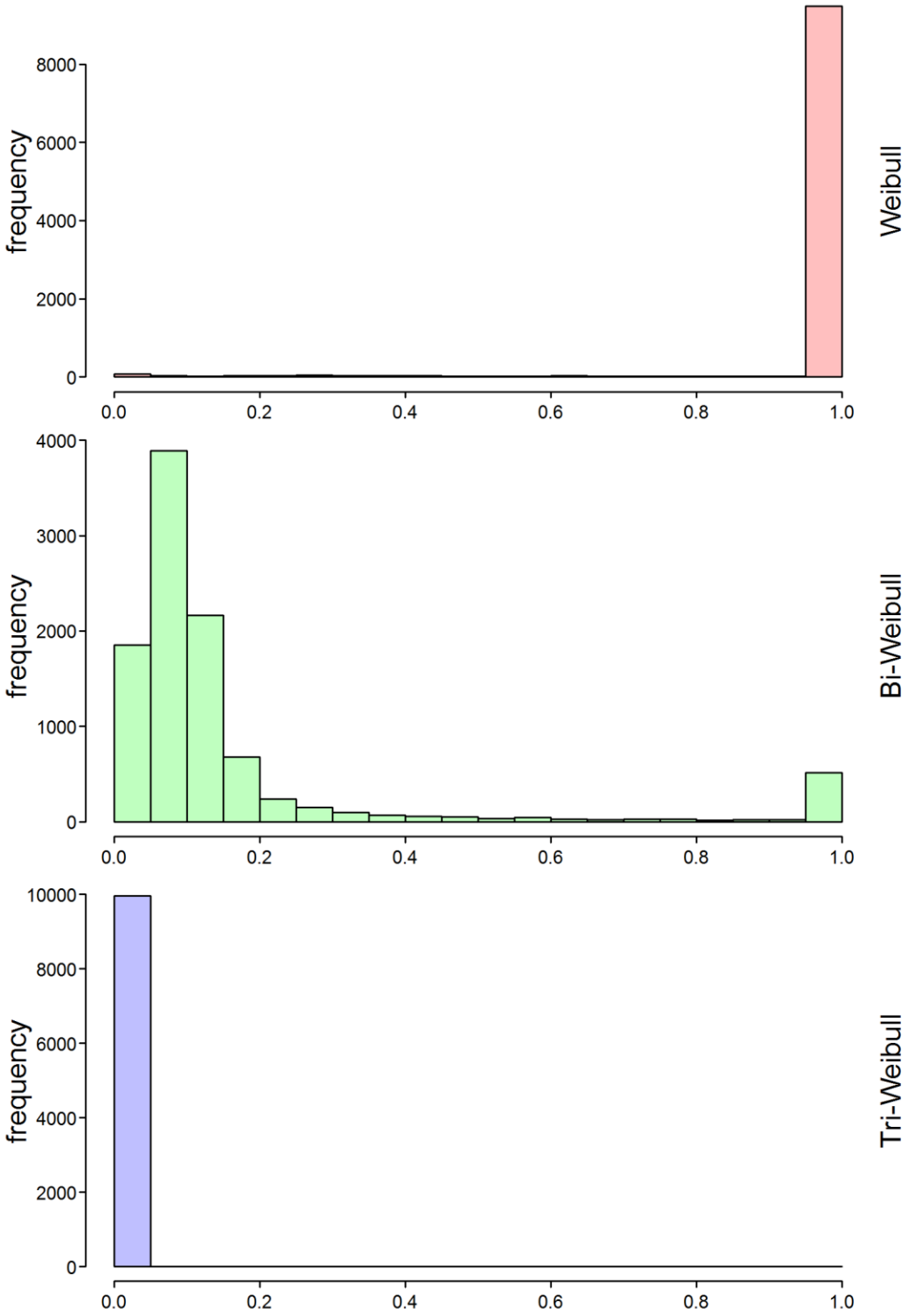

Figure 63 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 10$ systems, quality = high)

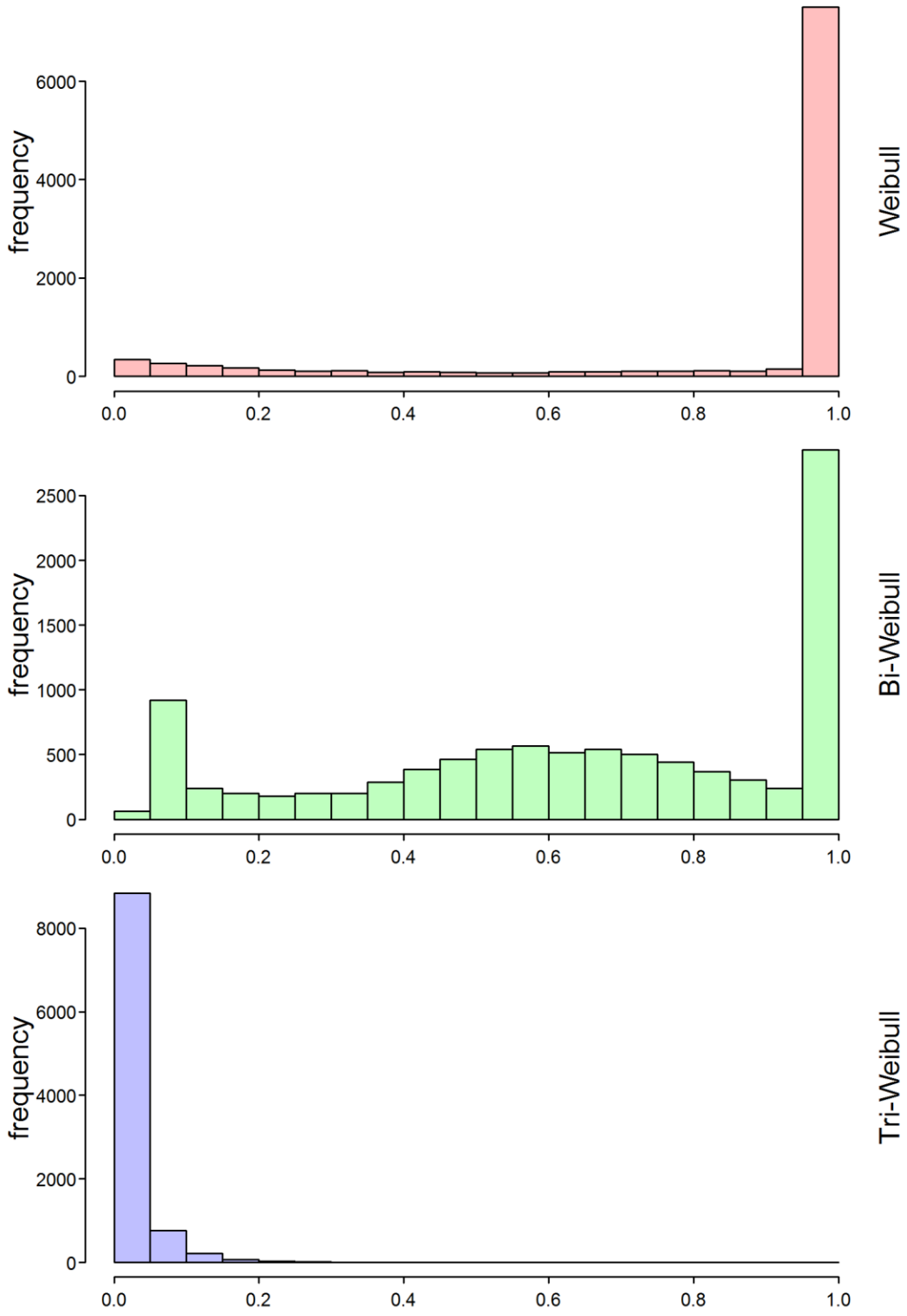

Figure 64 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 20$ systems, quality $=$ high)

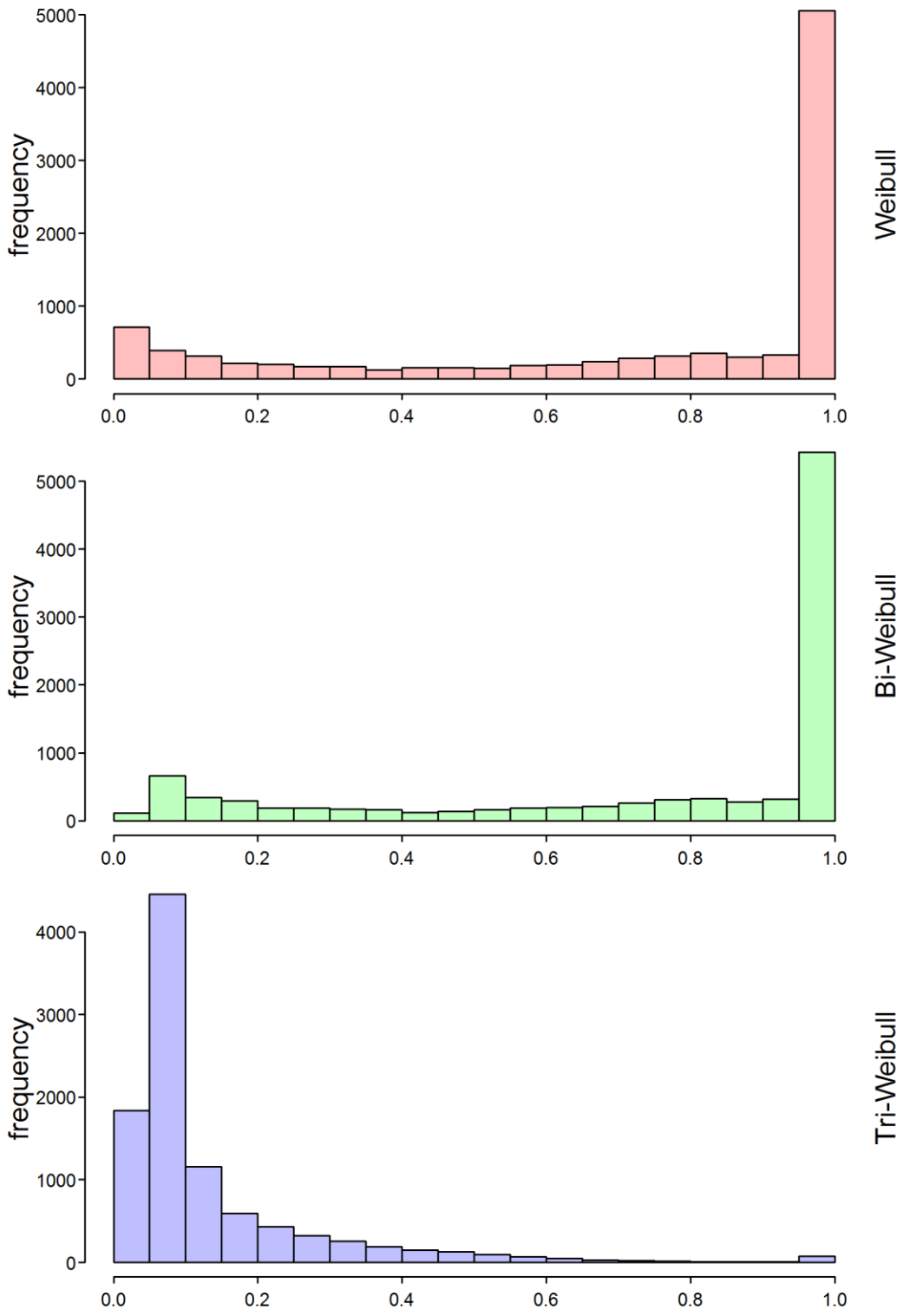

Figure 65 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 40$ systems, quality $=$ high)

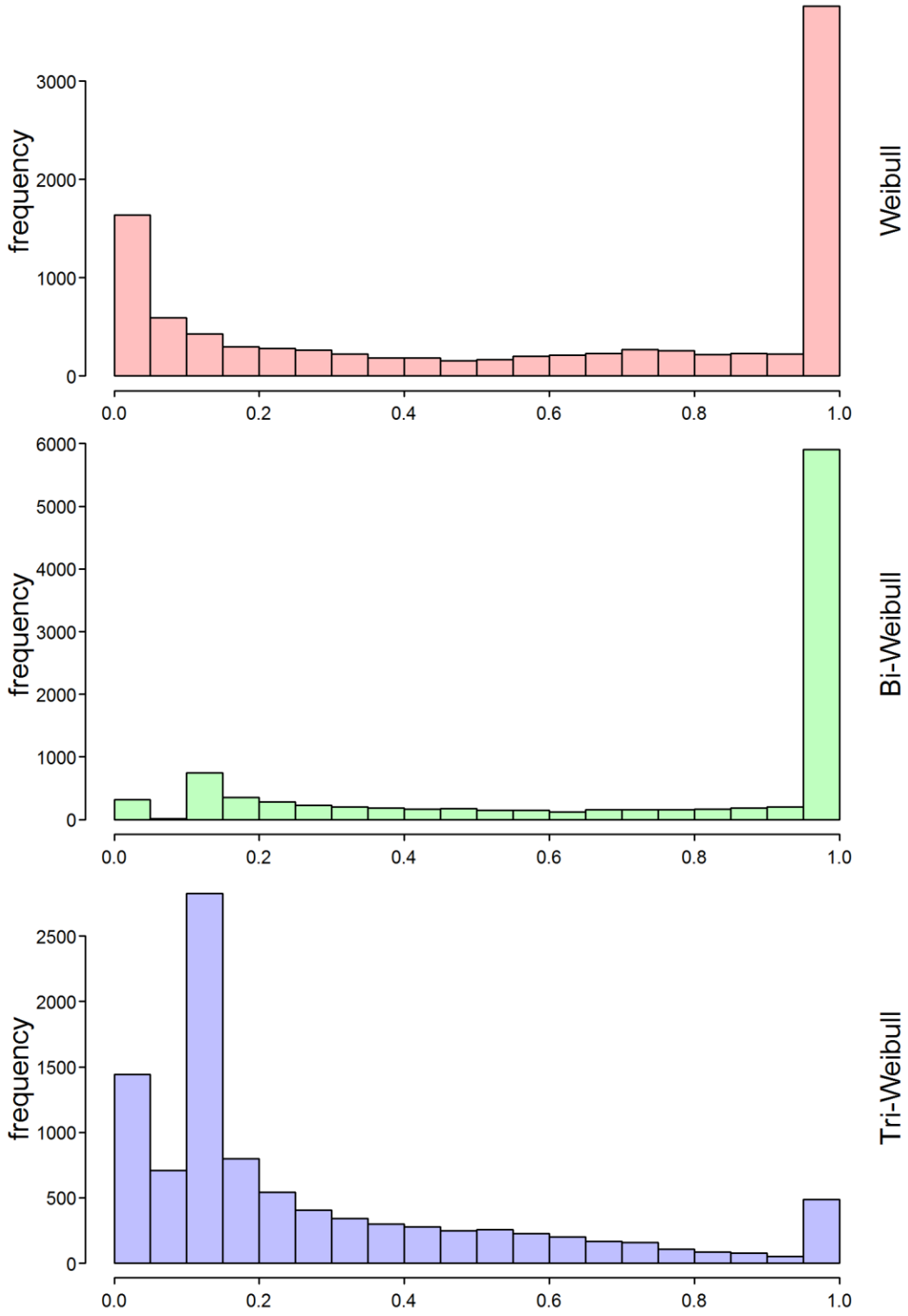

Figure 66 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 20$ modes, $n = 100$ systems, quality $=$ high)

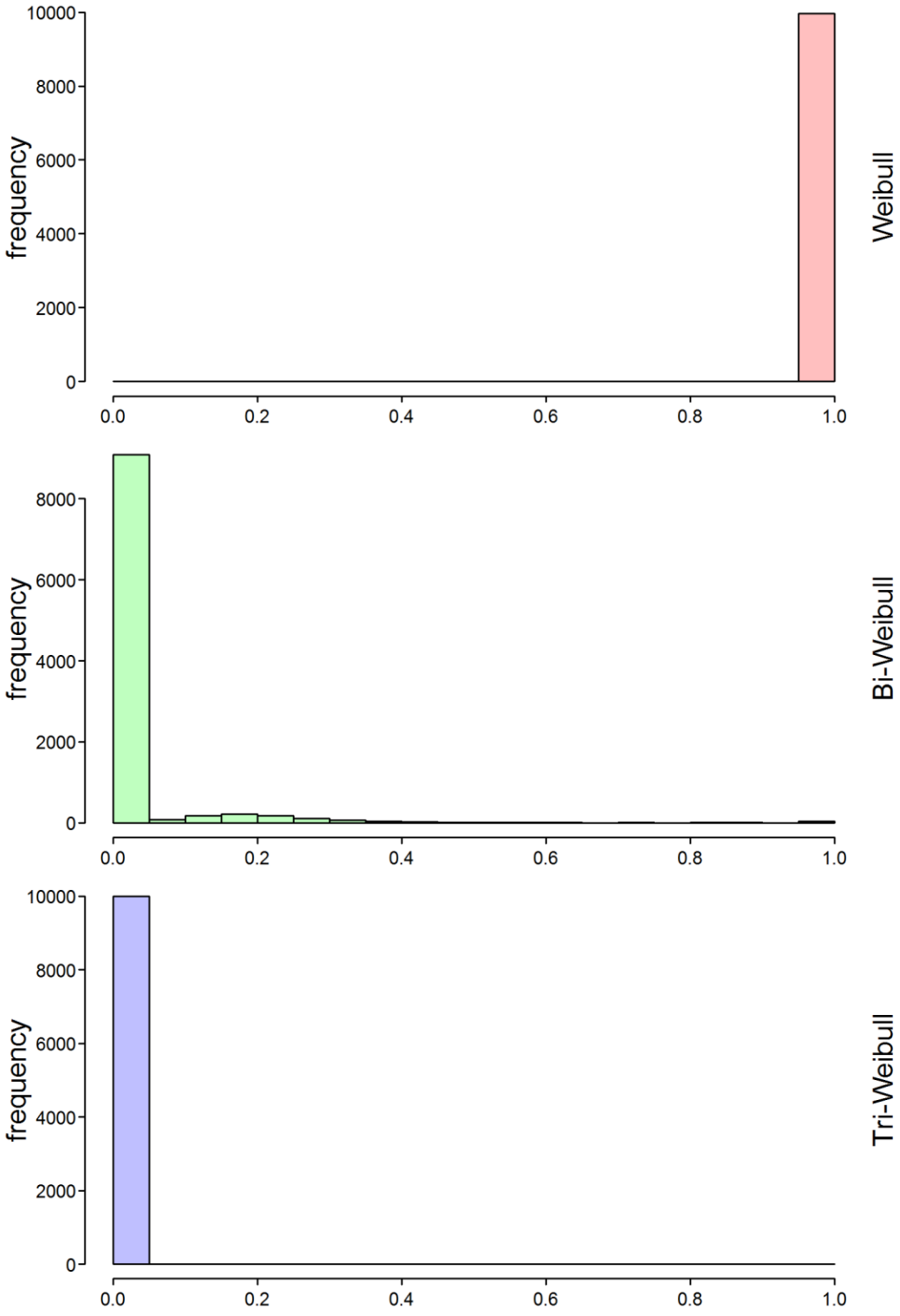

Figure 67 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 10$ systems, quality $=$ low)

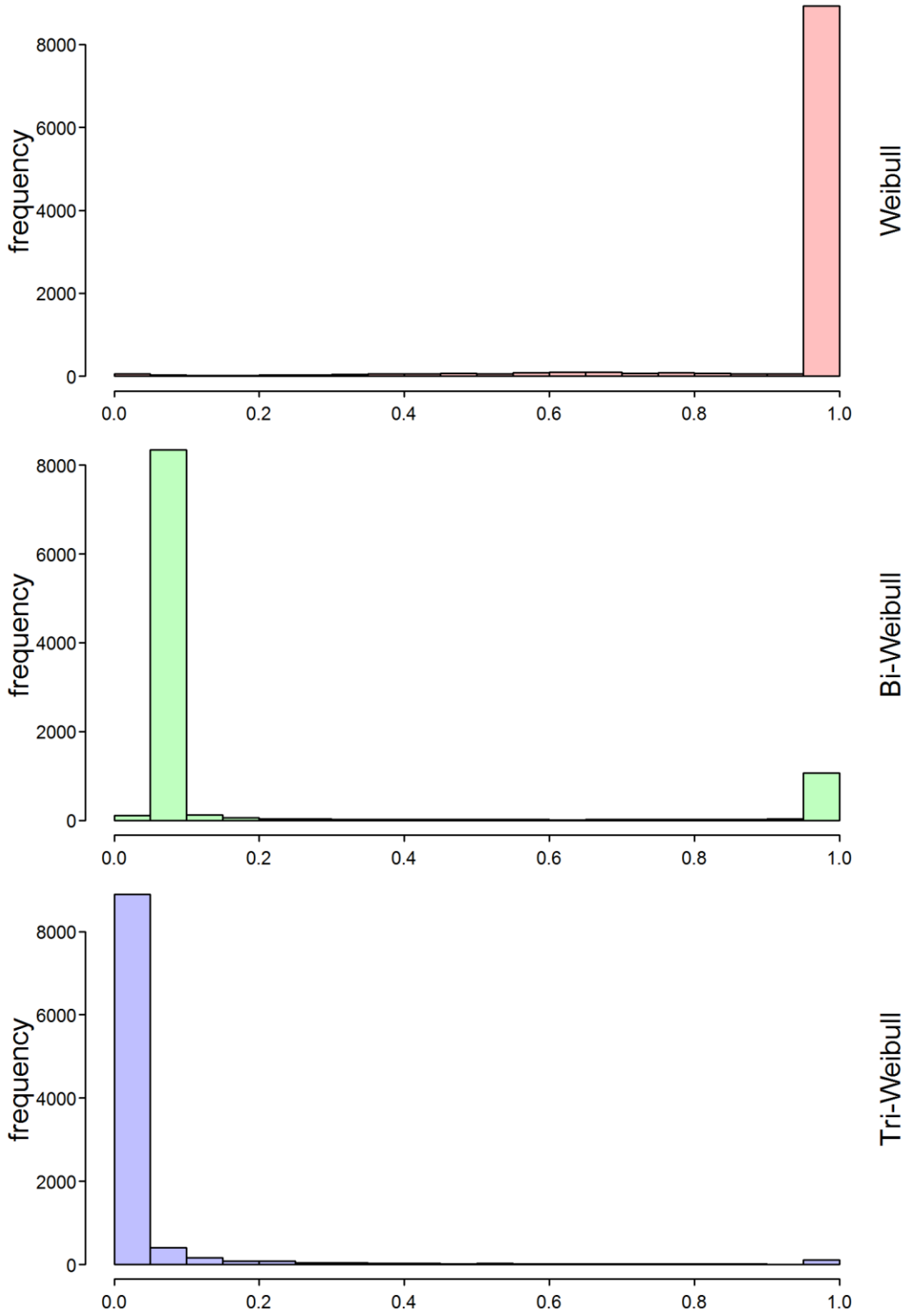

Figure 68 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 20$ systems, quality $=$ low)

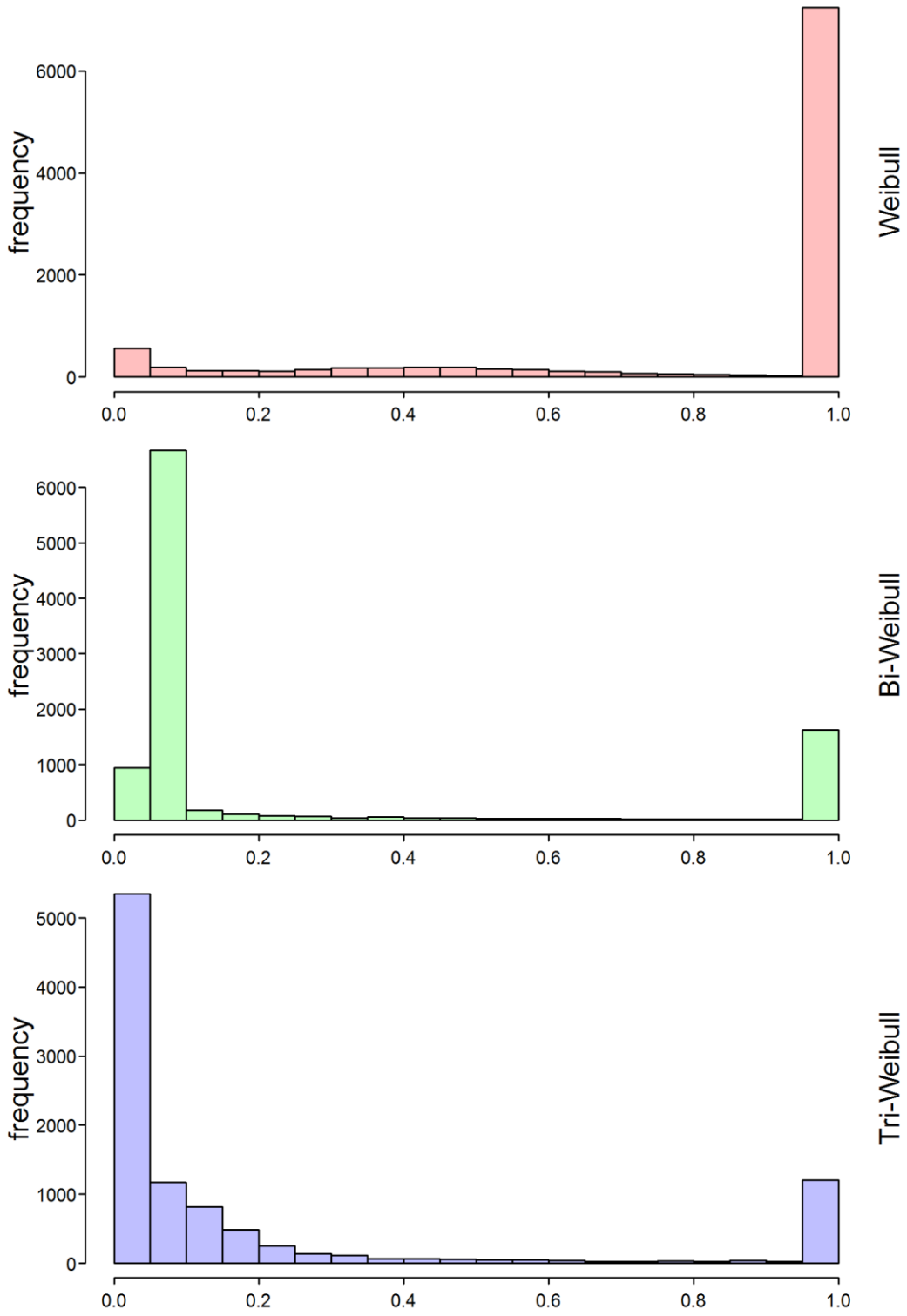

Figure 69 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 40$ systems, quality $=$ low)

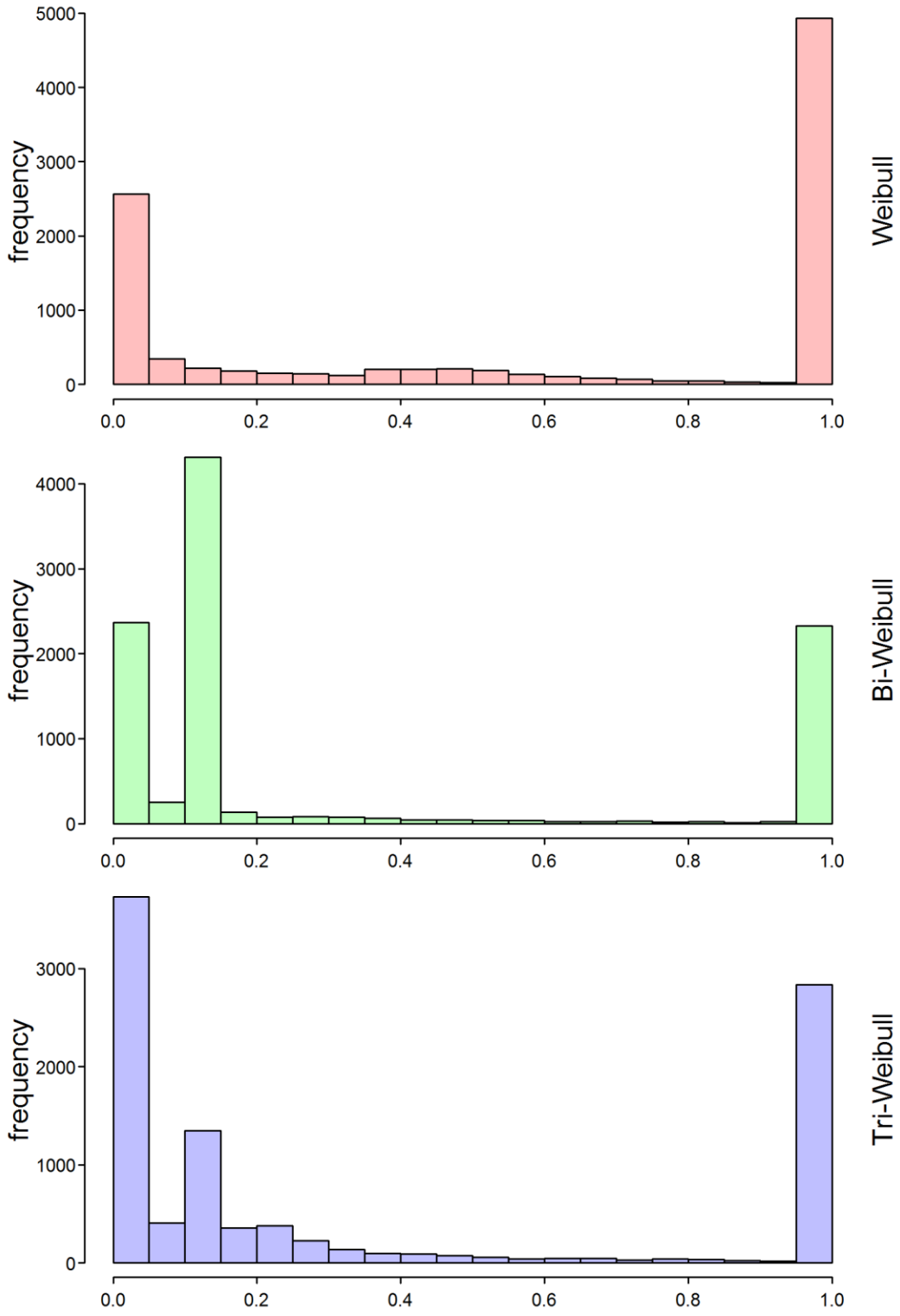

Figure 70 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 100$ systems, quality $=$ low)

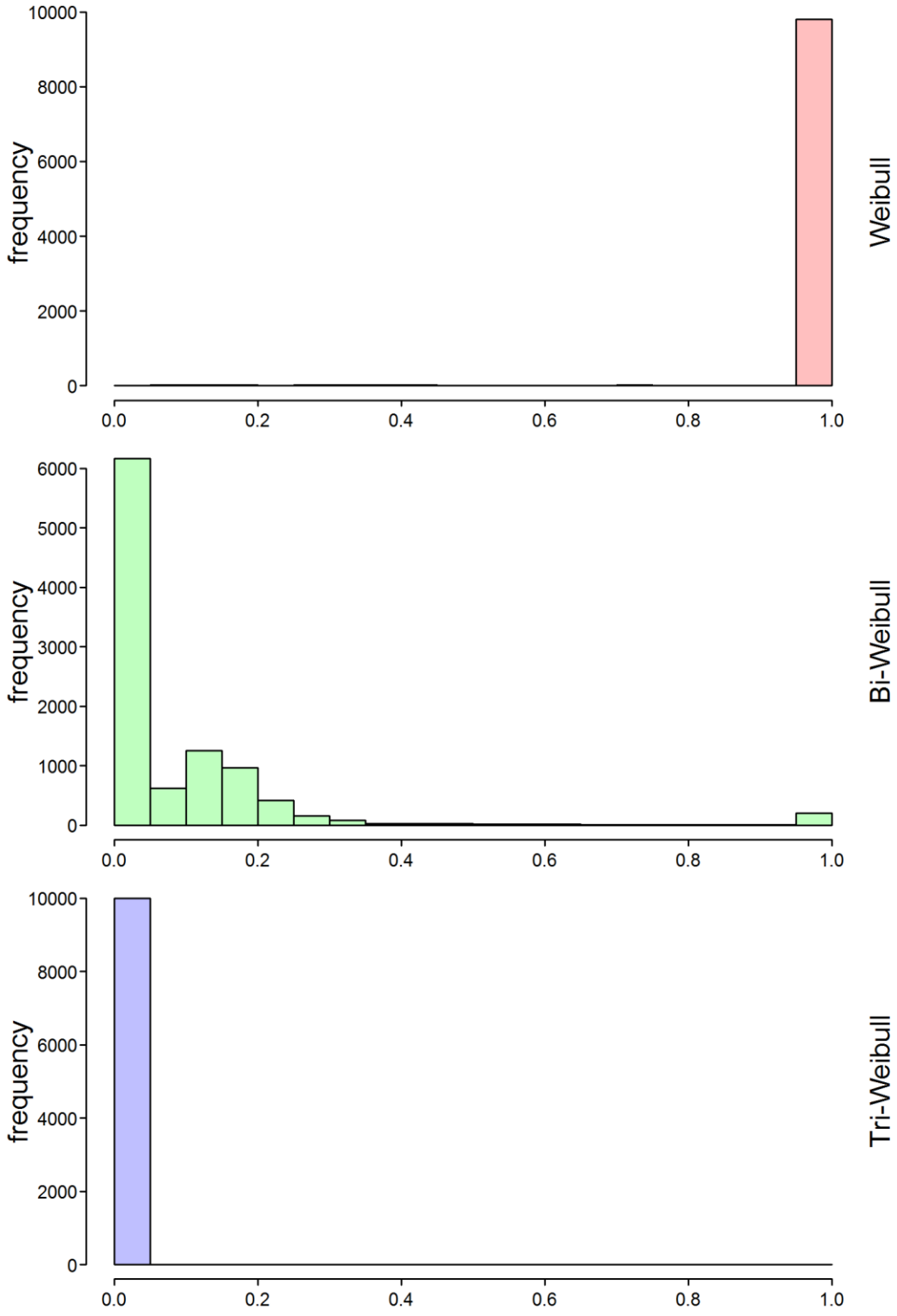

Figure 71 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 10$ systems, quality $=$ mid)

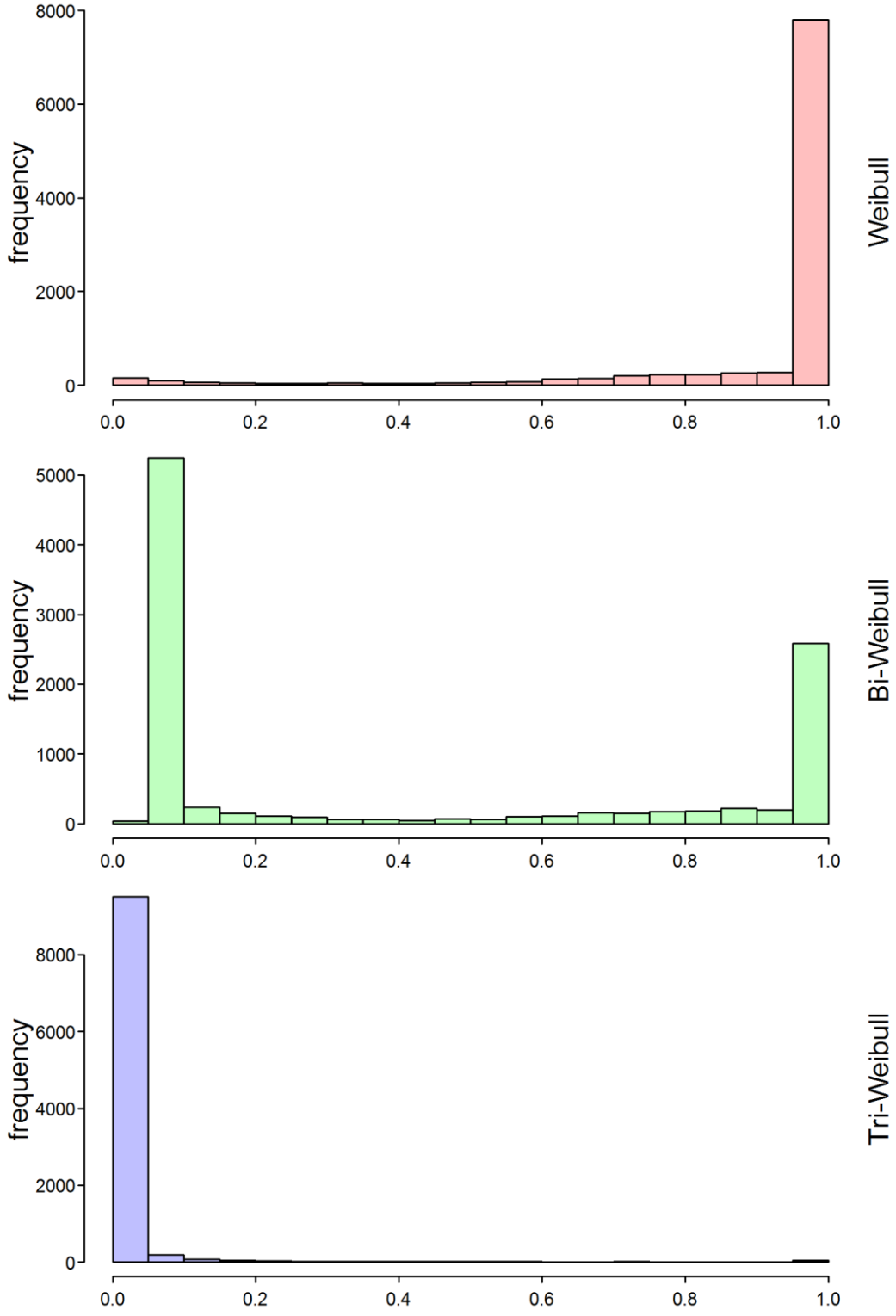

Figure 72 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 20$ systems, quality $=$ mid)

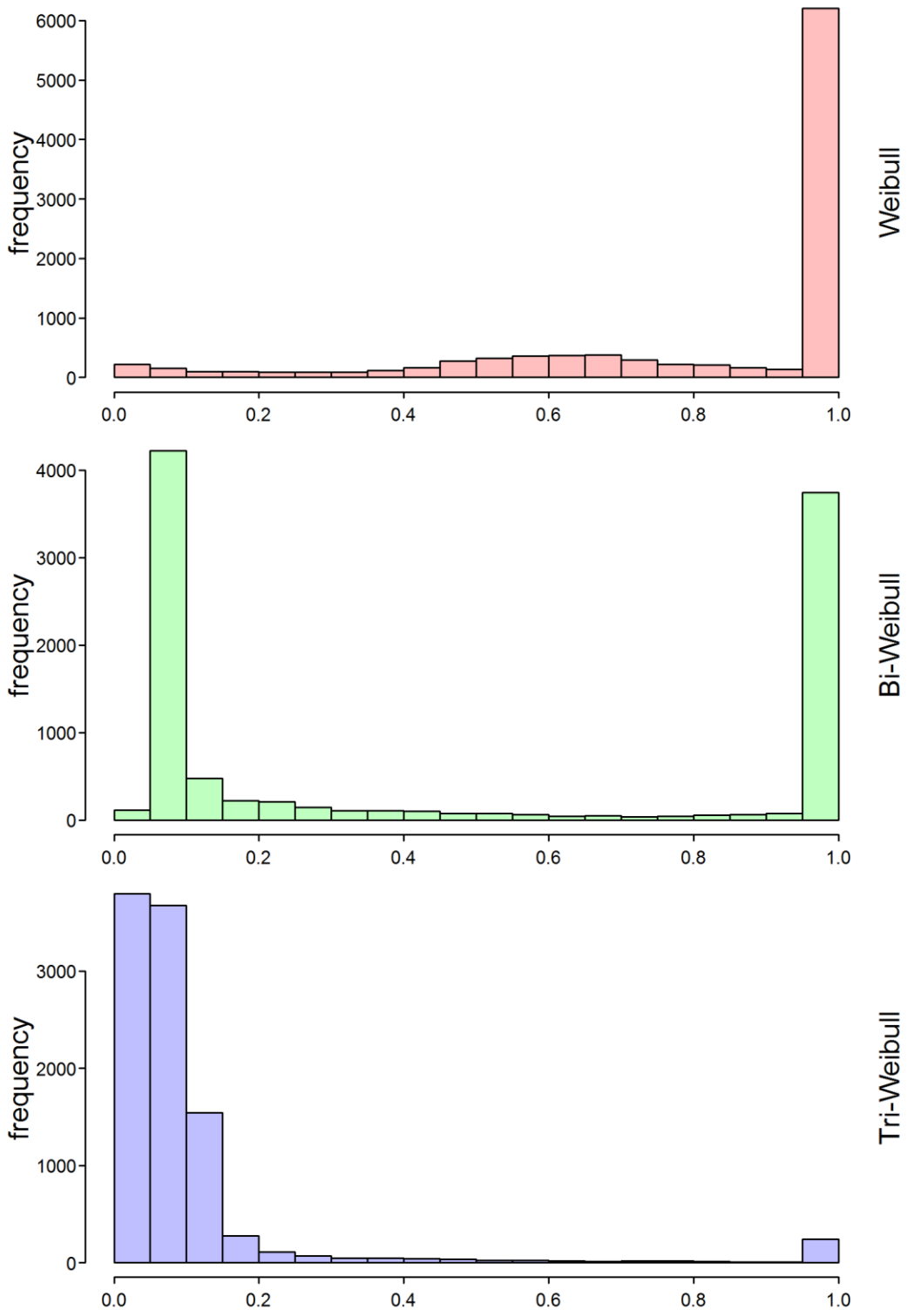

Figure 73 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 40$ systems, quality $=$ mid)

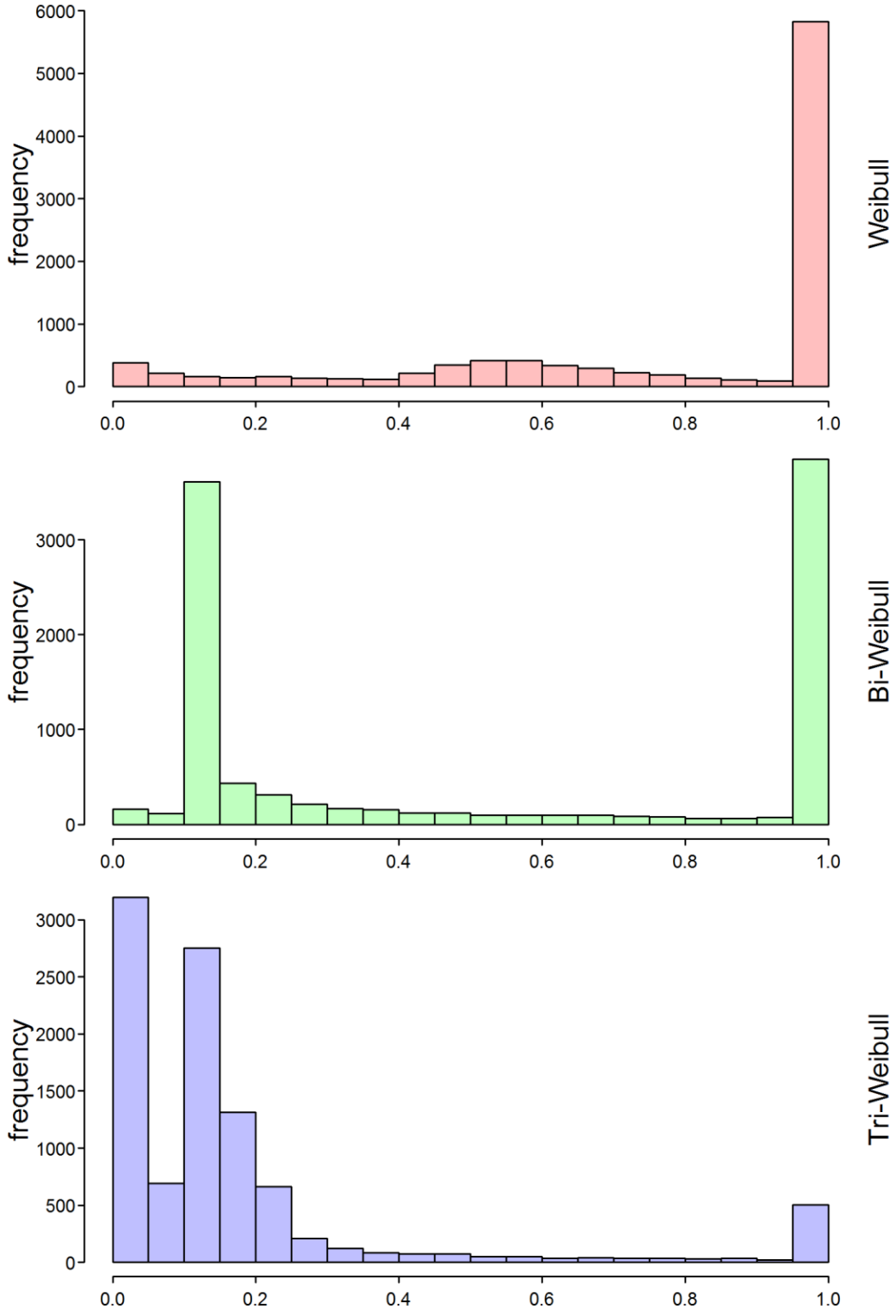

Figure 74 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 100$ systems, quality $=$ mid)

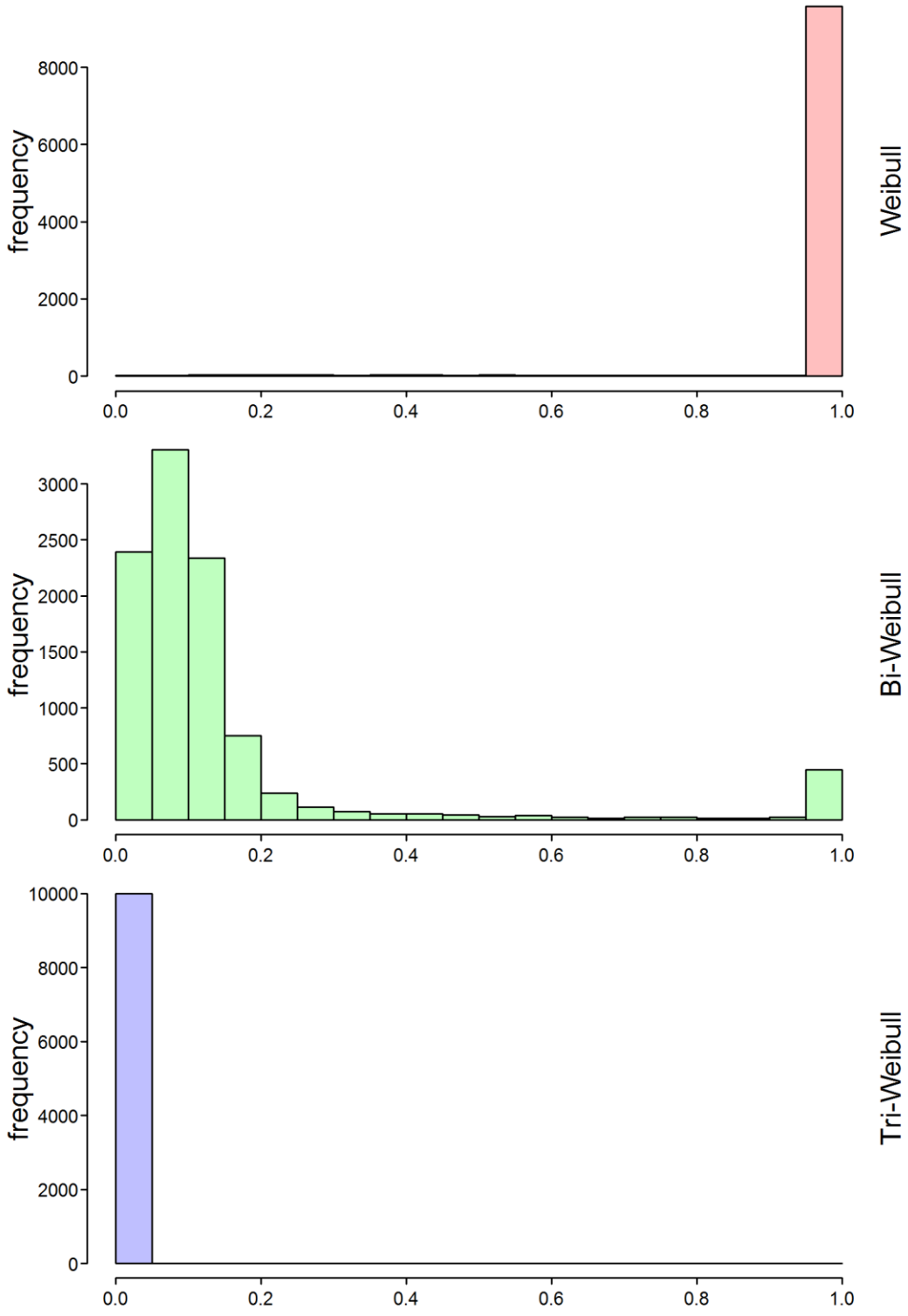

Figure 75 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 10$ systems, quality = high)

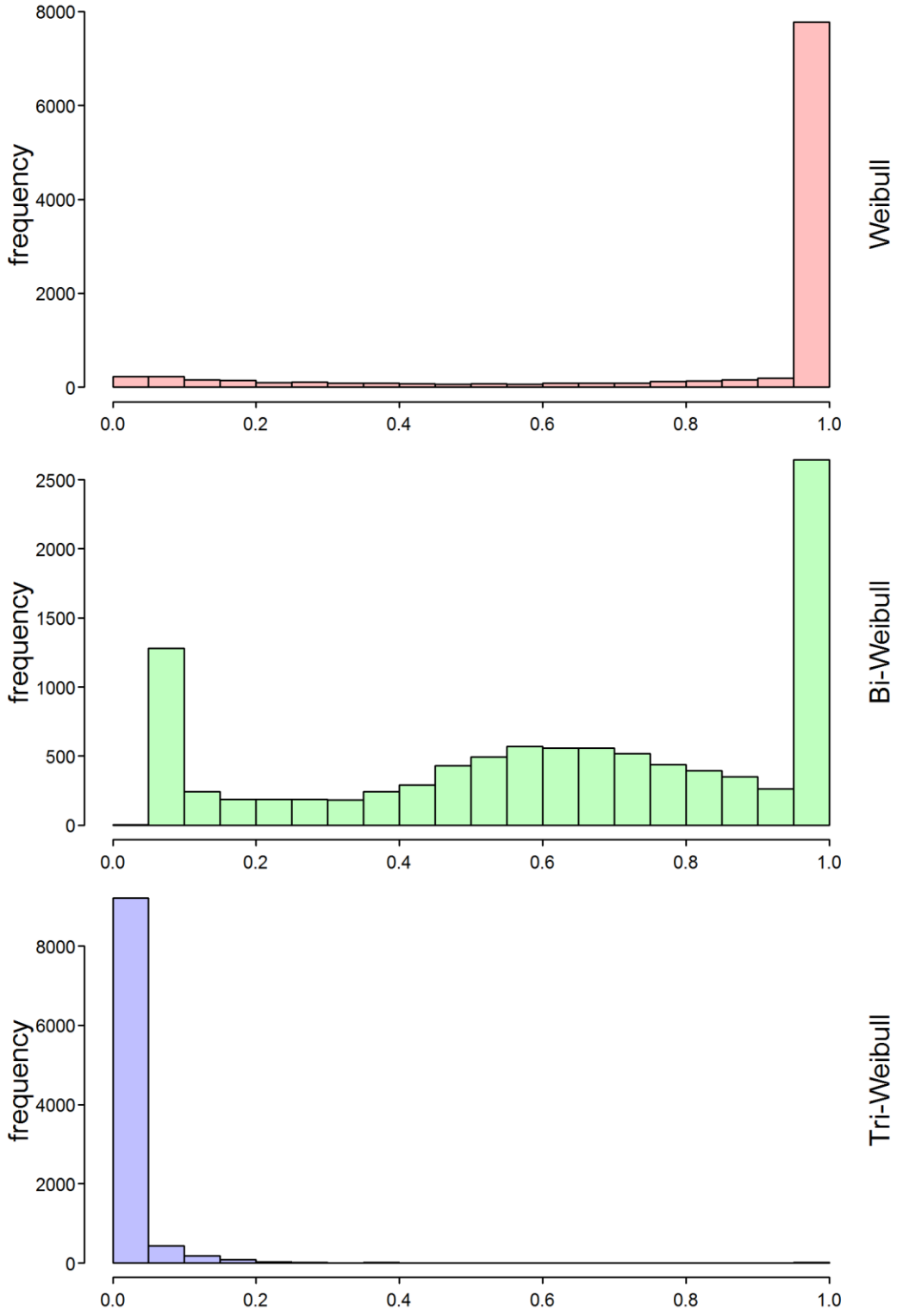

Figure 76 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 20$ systems, quality = high)

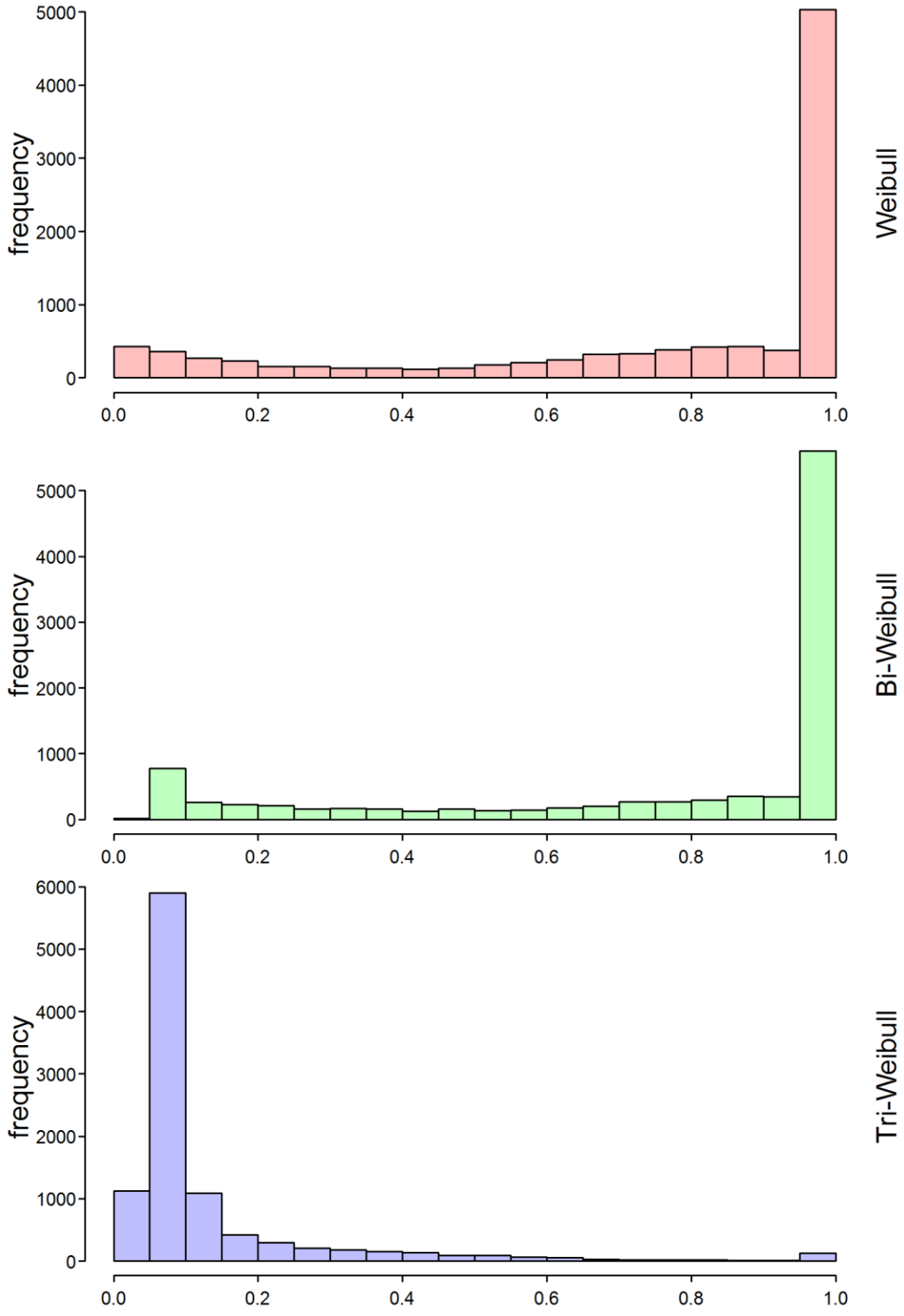

Figure 77 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 40$ systems, quality $=$ high)

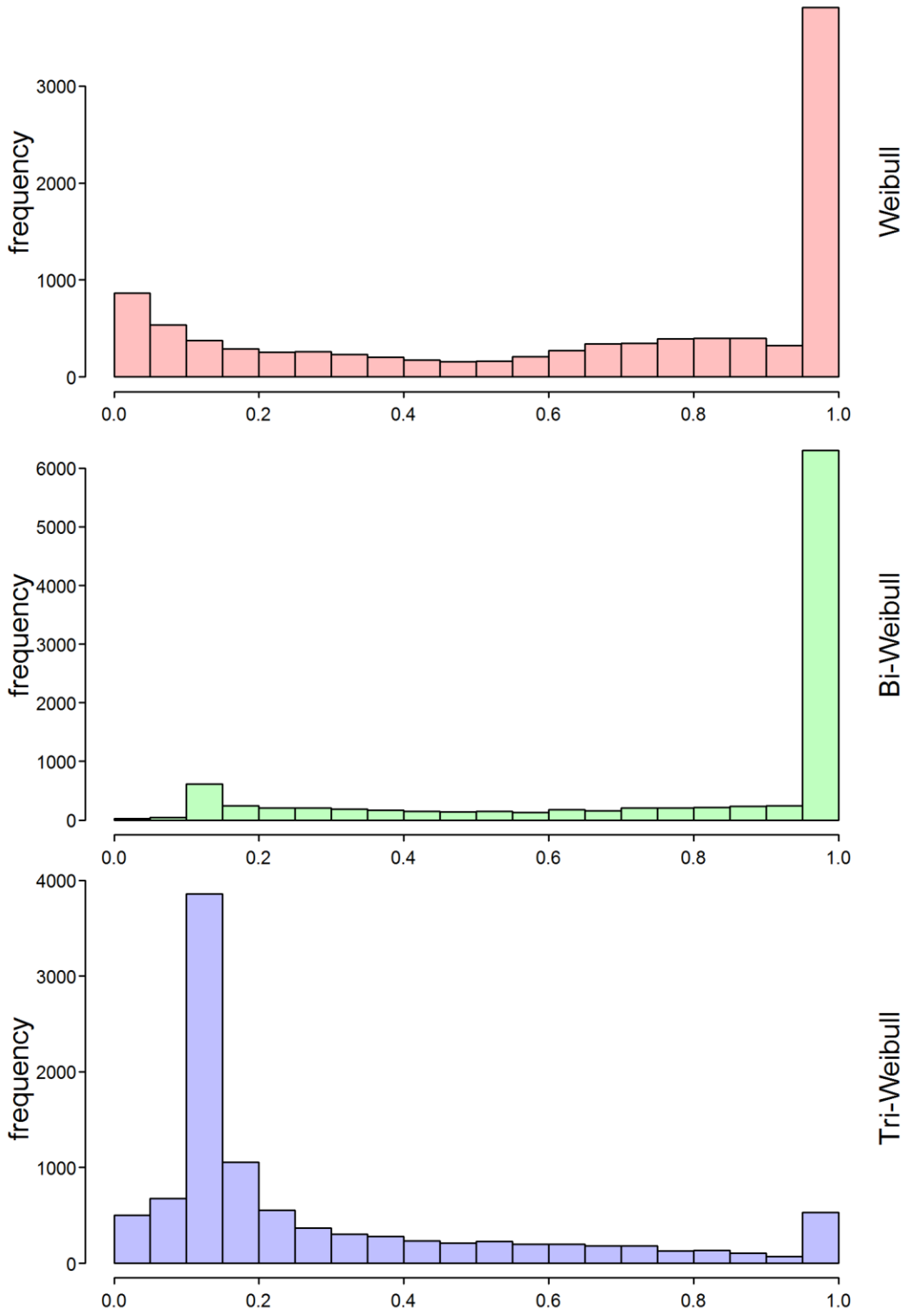

Figure 78 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 40$ modes, $n = 100$ systems, quality = high)

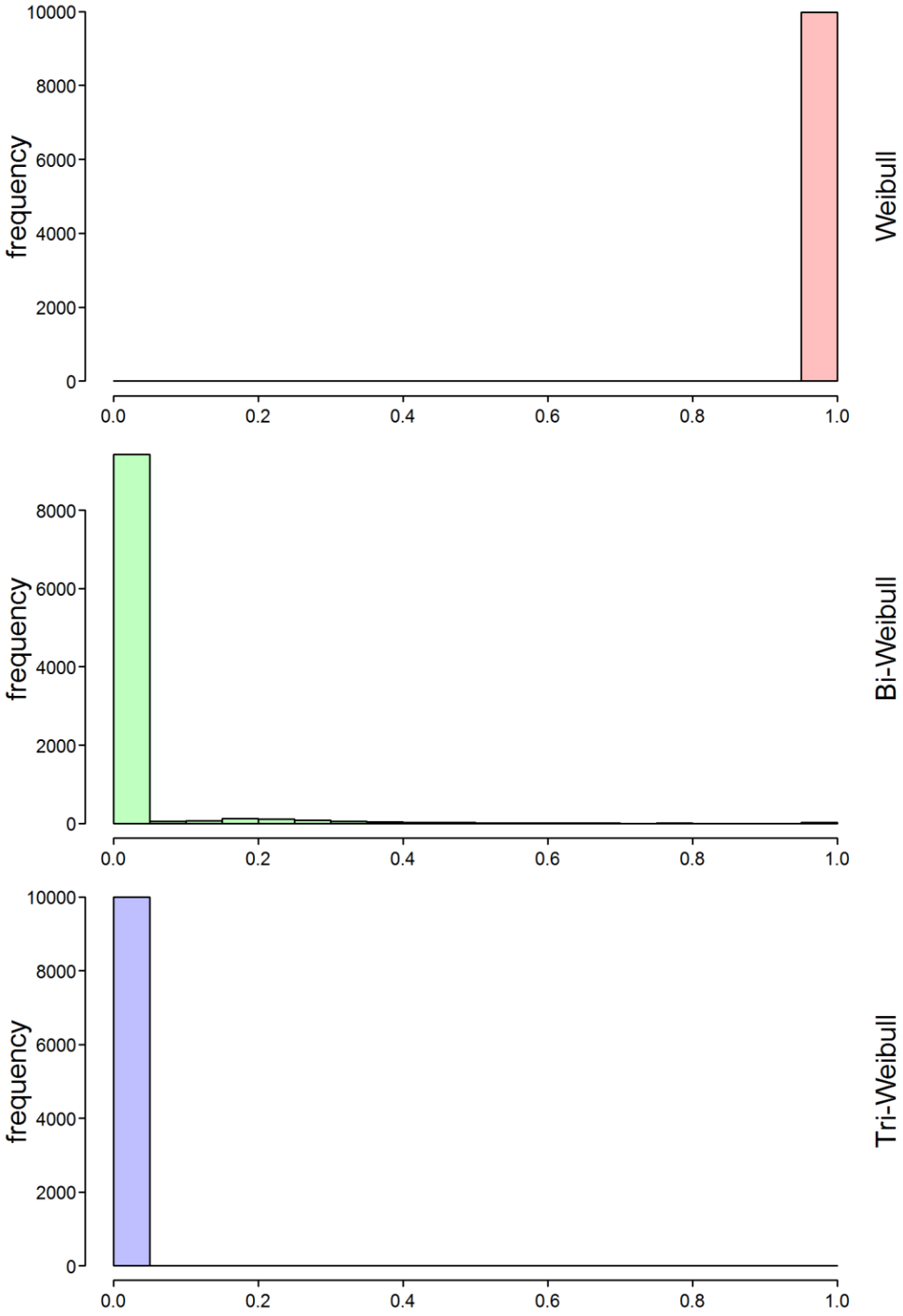

Figure 79 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 10$ systems, quality $=$ low)

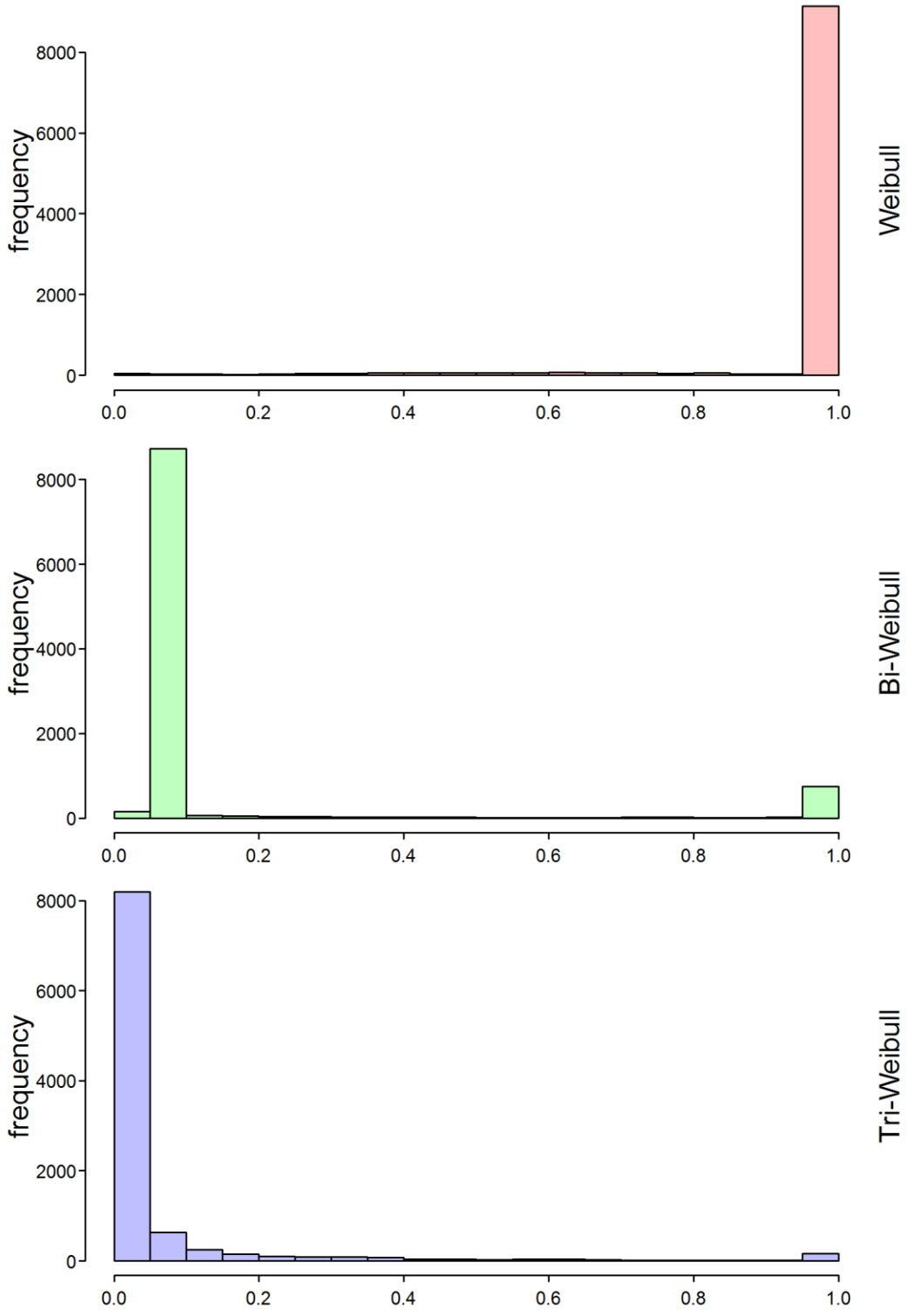

Figure 80 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 20$ systems, quality $=$ low)

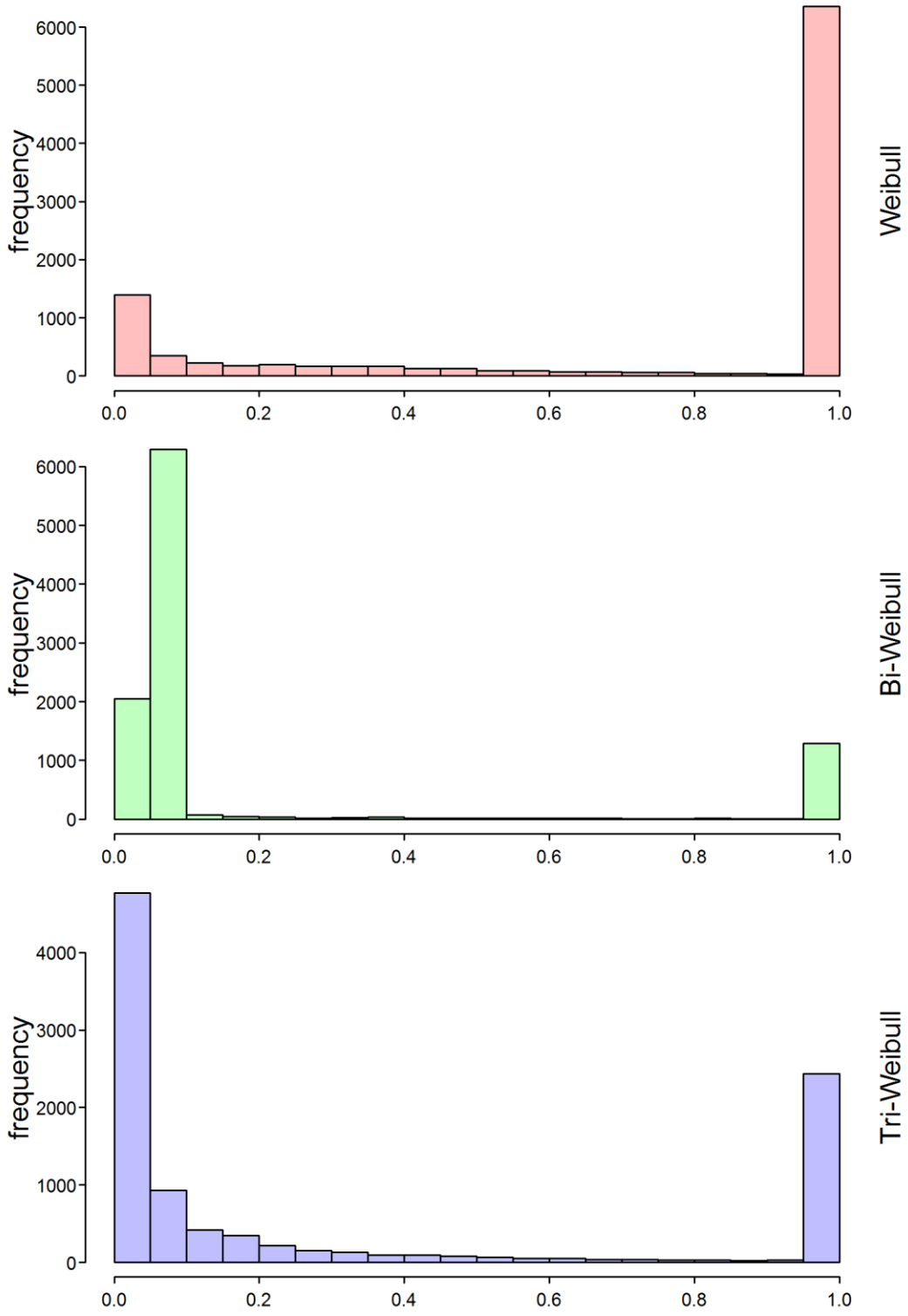

Figure 81 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 40$ systems, quality $=$ low)

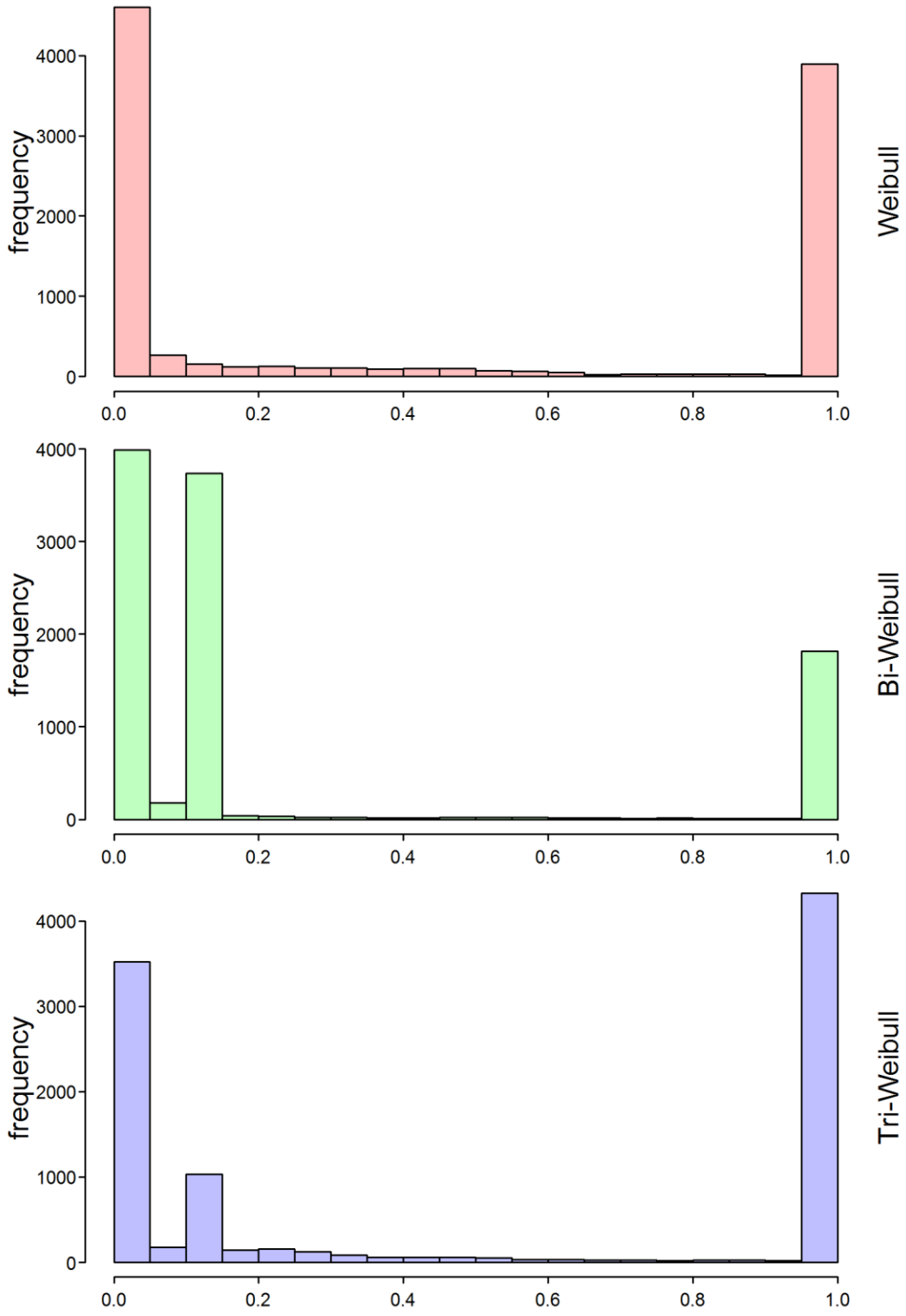

Figure 82 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 100$ systems, quality $=$ low)

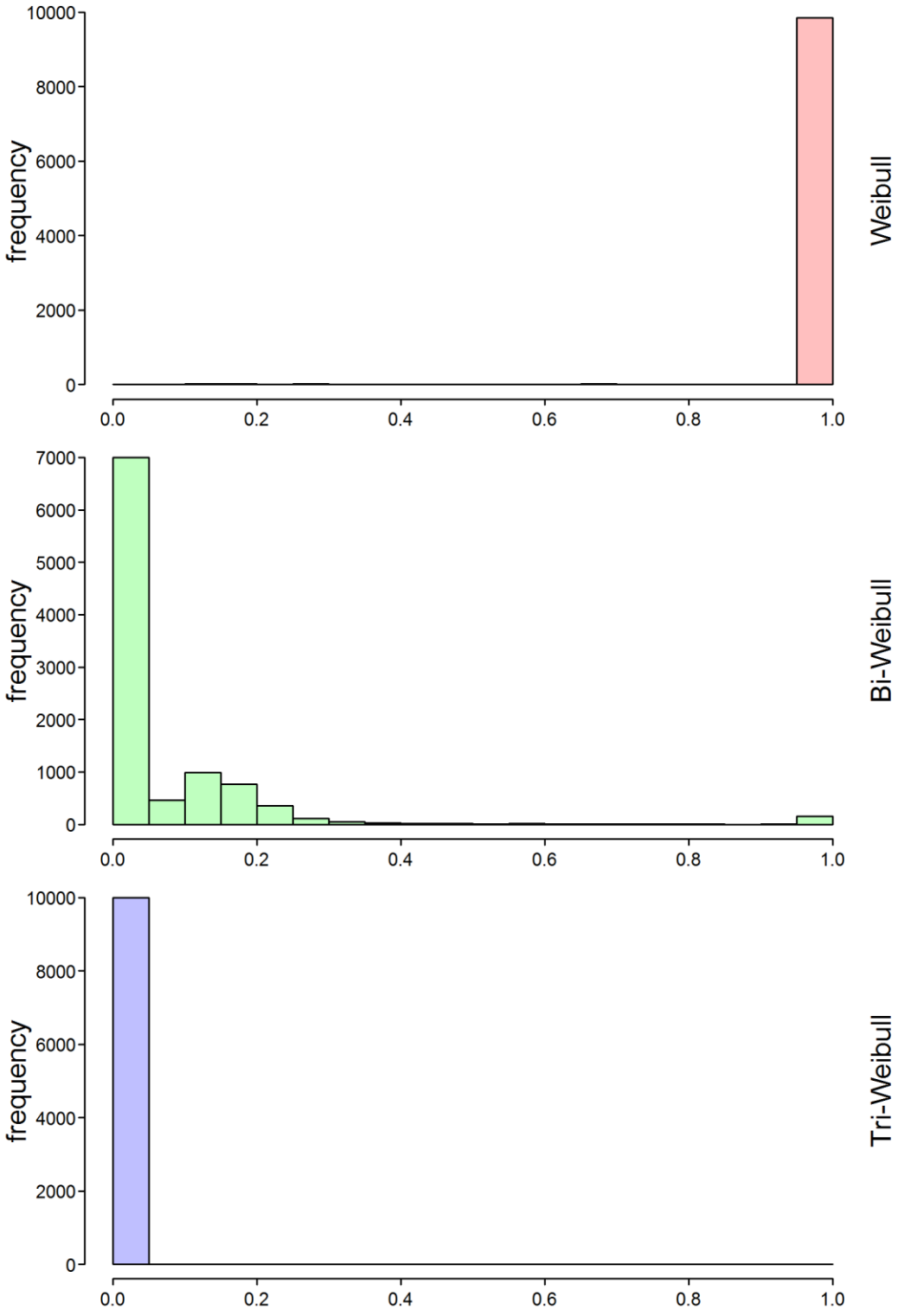

Figure 83 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 10$ systems, quality $=$ mid)

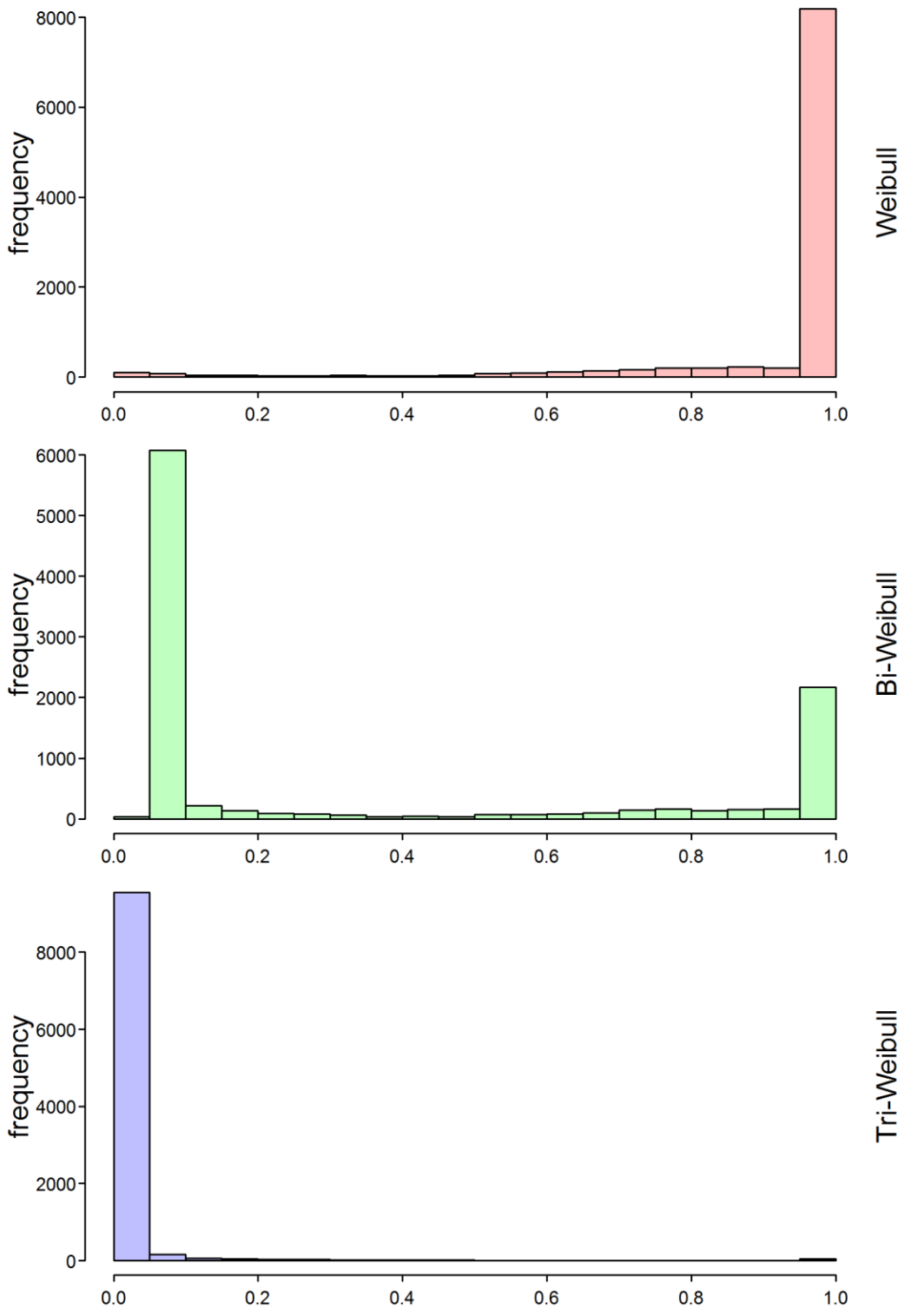

Figure 84 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 20$ systems, quality = mid)

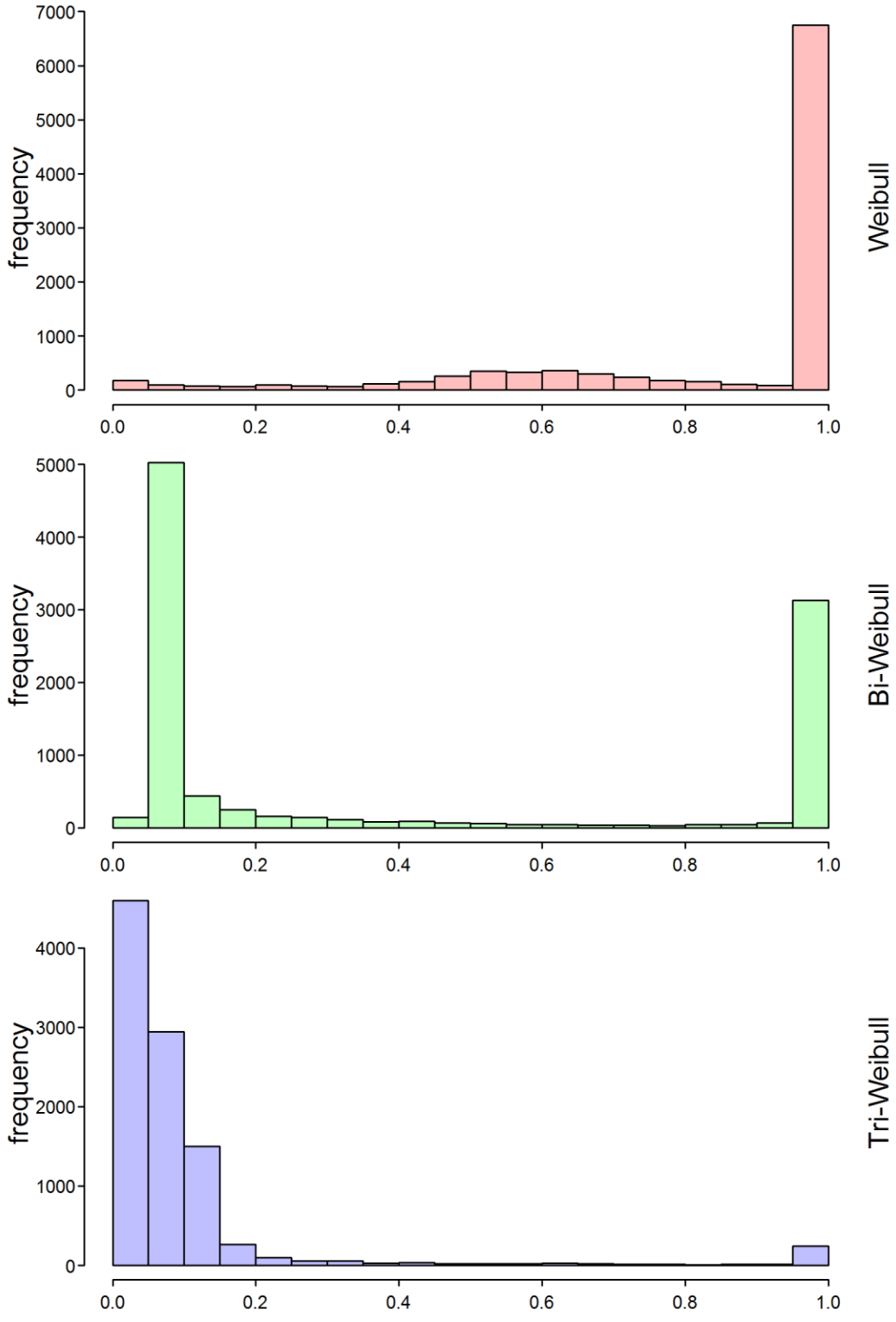

Figure 85 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 40$ systems, quality $=$ mid)

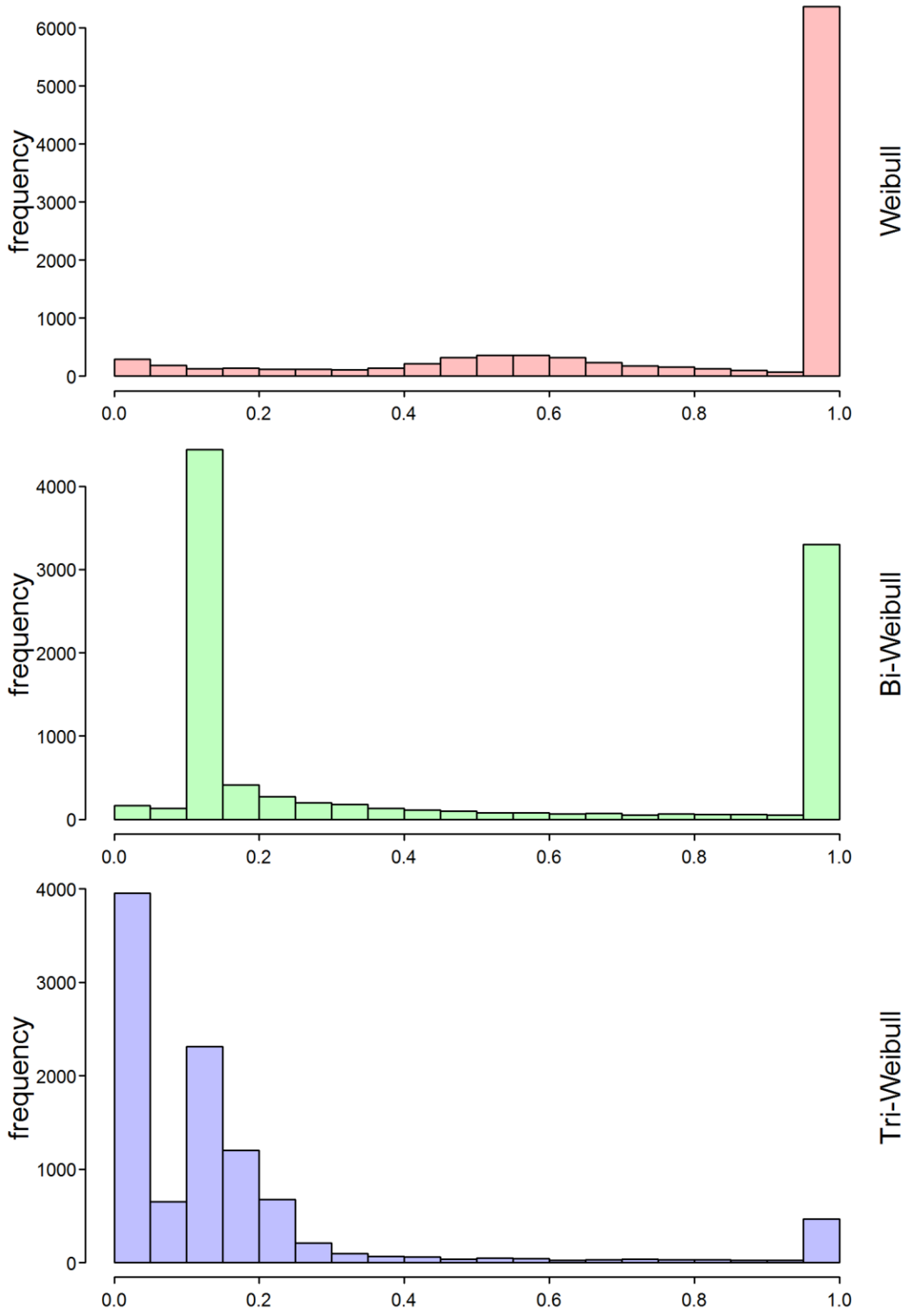

Figure 86 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 100$ systems, quality $=$ mid)

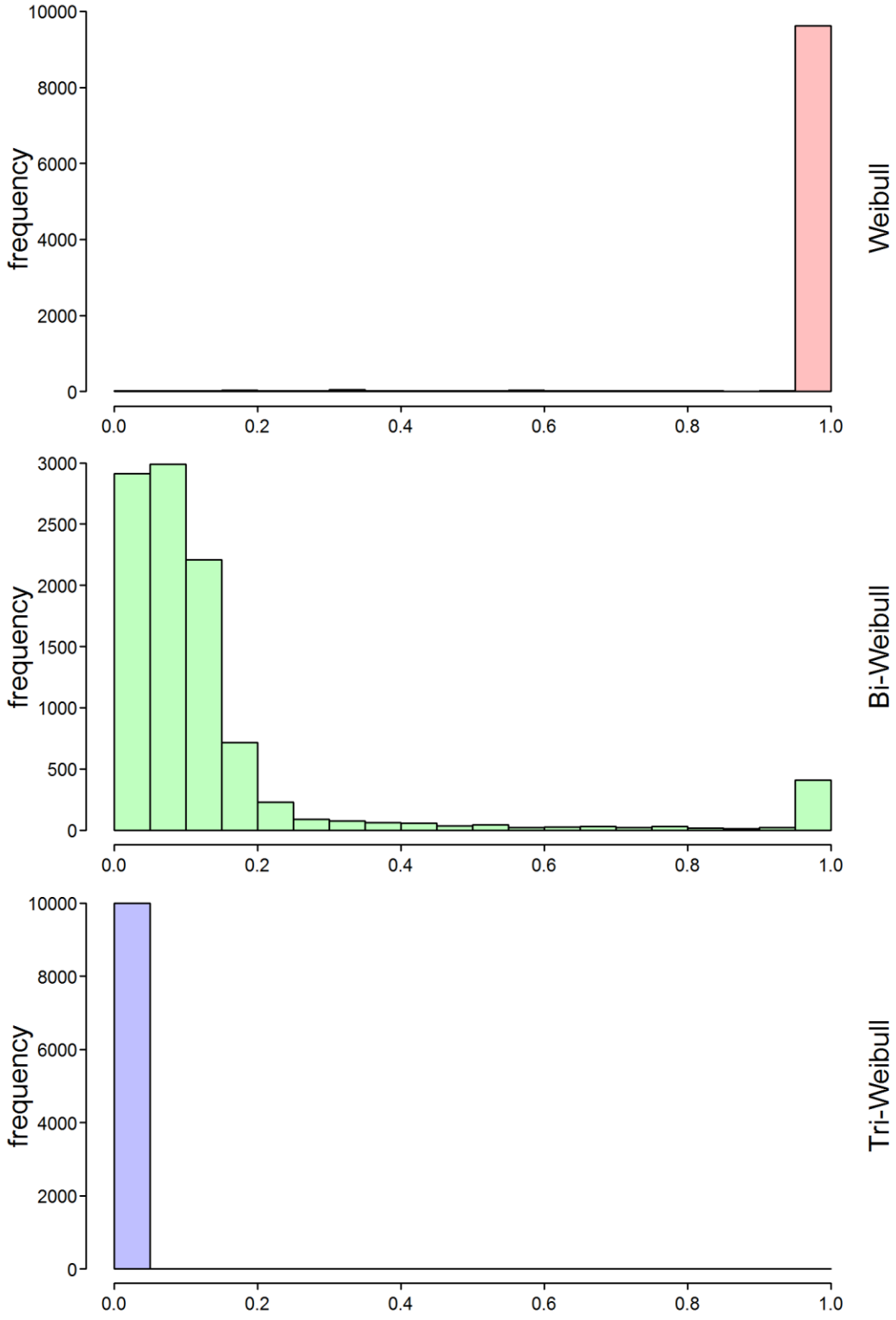

Figure 87 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 10$ systems, quality $=$ high)

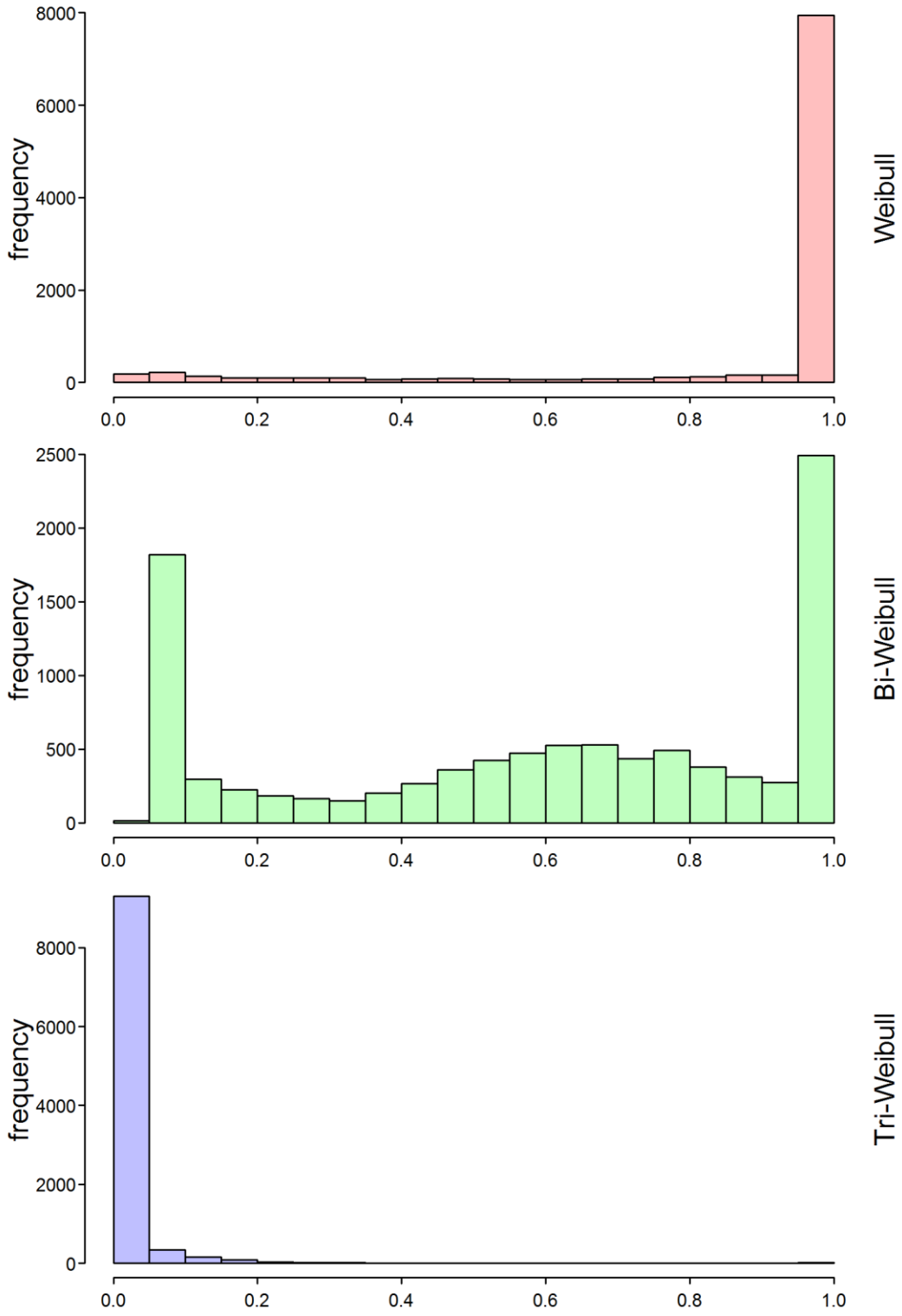

Figure 88 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 20$ systems, quality $=$ high)

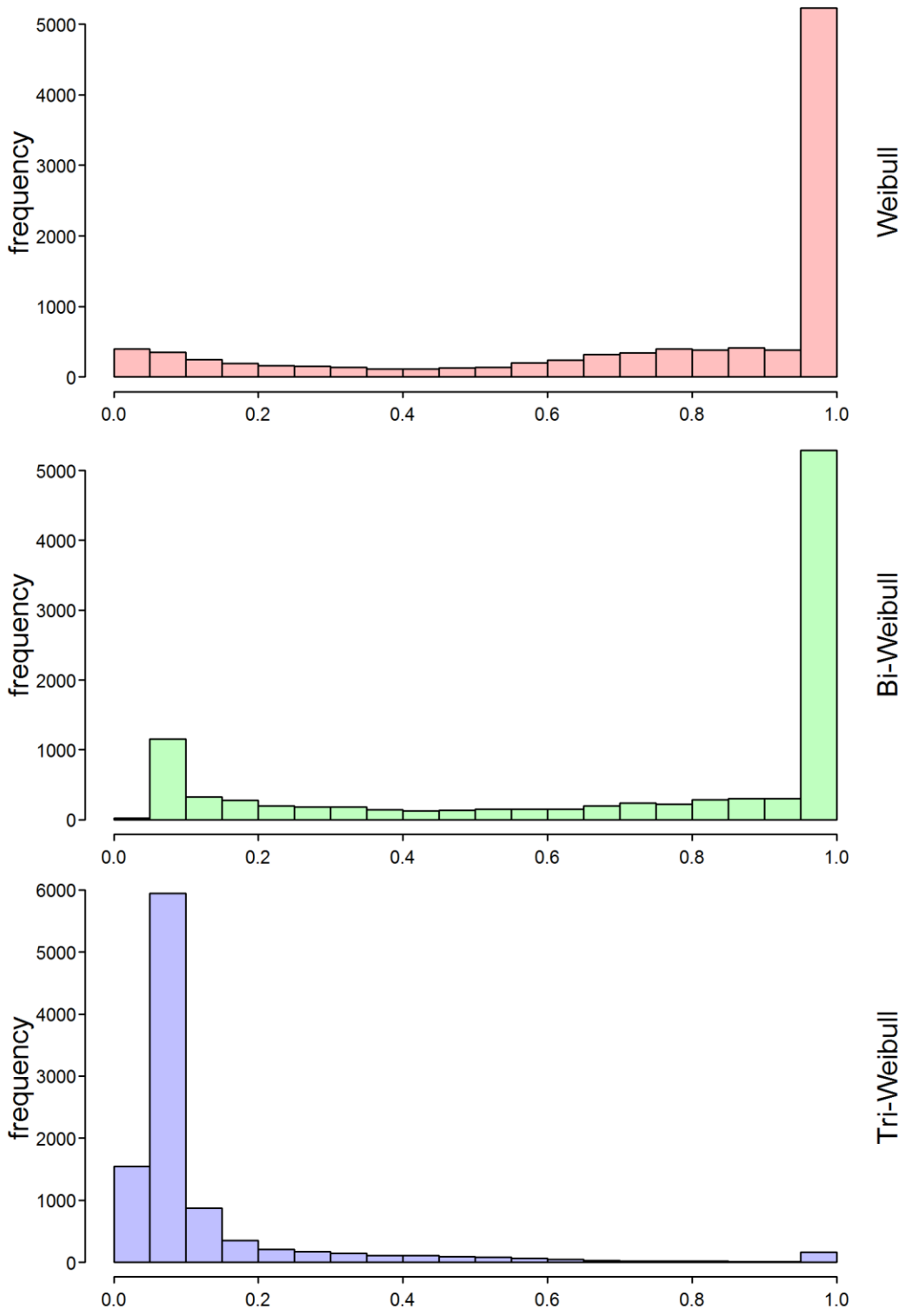

Figure 89 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 40$ systems, quality = high)

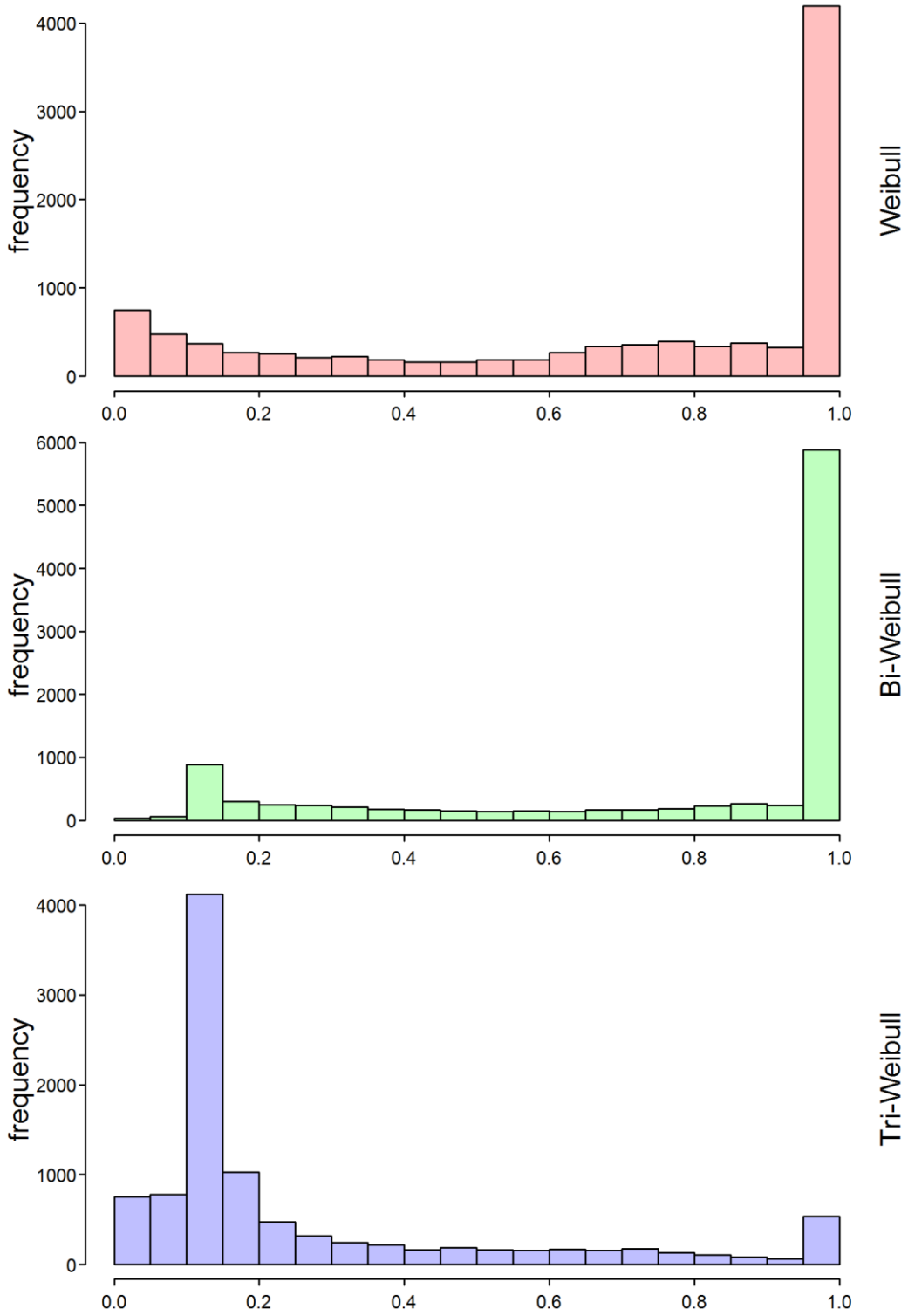

Figure 90 – Histogram of $AICc$ values for the Weibull, bi-Weibull and tri-Weibull distributions ($J = 50$ modes, $n = 100$ systems, quality $=$ high)

Appendix C: Computer Code Used to Compare the Fit of Various Distributions to the Aarset and Meeker Datasets as Discussed in Chapter IV

```
##### Meeker Data set
Meeker<-
sort(c(275,13,147,22.5,23.5,181,30,65,10,173,106,212,2,261,293,88,247,28,143,80,
245,266,300,300,300,300,300,300,300,300))
Meeker1<-
sort(c(275,13,147,22.5,23.5,181,30,65,10,173,106,212,2,261,293,88,247,28,143,80,
245,266))#300,300,300,300,300,300,300,300))
Meeker2<-c(300,300,300,300,300,300,300,300)

Censor<-c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0,0,0)
eCDF<-
sort(c(275,13,147,22.5,23.5,181,30,65,10,173,106,212,2,261,293,88,247,28,143,80,
245,266,300))
RelFreq<-
c(0.03333,.06667,0.1,.13333,.16667,.2,.2333,.26667,.3,.3333,.36667,.4,.43333,.4666
7,.5,.53333,.56667,.6,.63333,.66667,.7,.73333,1)

##### Aarset Data set
Aarset<-
c(0.1,0.2,1,1,1,1,1,2,3,6,7,11,12,18,18,18,18,18,21,32,36,40,45,46,47,50,55,60,63,63,
67,67,67,67,72,75,79,82,82,83,84,84,84,85,85,85,85,85,86,86)
Aarset2<-c(Aarset+rnorm(50, 0, 0.001))
RelFreq<-c(rep(1/length(Aarset2),length(Aarset2)))*c(1:length(Aarset2))

Red<-c(0.1, 0.2,
1,2,3,6,7,11,12,18,21,32,36,40,45,46,47,50,55,60,63,67,72,75,79,82,83,84,85,86)
RedAA<-
c(.02,.04,.14,.16,.18,.2,.22,.24,.26,.36,.38,.4,.42,.44,.46,.48,.5,.52,.54,.56,.6,.68,.7,.72,.74,
.78,.8,.86,.96,1)

######### Compare Fit of Poly-Weibull/New Modified Weibull Reliability Against
Meeker Data ############

######## New Modified Weibull Model ##########################
B0<-matrix(NA, nrow=100, ncol=5, byrow=FALSE,); colnames(B0)<-
c("a","b","A","g","l")
B0[,1]<-c(runif(100, 0.01,1)) #a
B0[,2]<-c(runif(100, 0.01,1)) #b
B0[,3]<-c(runif(100, 0.01,1)) #t
B0[,4]<-c(runif(100, 0.01,1)) #g
B0[,5]<-c(runif(100, 0.01,1)) #l
```

```
NMW<-function(x,y,d){
 a<-x[1]
 b<-x[2]
 t<-x[3]
 g<-x[4]
 l<-x[5]

 n<-length(Meeker)

 F<-c(rep(NA,1))

 F<-(sum(log(b*(g+l*y)*(y^(g-1))*exp(l*y)+a*t*y^(t-1))*d)-a*sum(y^t)-
b*sum((y^g)*exp(l*y)))

 return(-F)
}

N<-matrix(NA, nrow=100, ncol=6, byrow=FALSE,); colnames(N)<-
c("a","b","t","g","l","Value")

for (i in 1:100){
 N[i,]=0
 N[i,1:5]<-optim(par=B0[i,], fn=NMW, y=c(Meeker),d=Censor, method="L-BFGS-B",
lower=c(0.0001, 10^-9, 0.0001, 0.0001, 0.0001), upper=c(1, 1, 1, 1, 1)) $par
 N[i,6]<-NMW(N[i,1:5],c(Meeker),d=Censor)
}

NM<-which.min(N[,6])

####New Modified Weibull PDF
NMWPdf<-function(t){(N[NM,1]*N[NM,3]*t^(N[NM,3]-
1)+(N[NM,2]*(N[NM,4]+N[NM,5]*t)*t^(N[NM,4]-1)*exp(N[NM,5]*t)))*(exp(-
1*(N[NM,1]*t^N[NM,3]+N[NM,2]*t^N[NM,4]*exp(N[NM,5]*t))))}
####New Modified Weibull Reliability function
NMWeib<-function(t){exp(-
1*(N[NM,1]*t^N[NM,3]+N[NM,2]*t^N[NM,4]*exp(N[NM,5]*t)))}
####New Modified Weibull Hazard Function
NMWHaz<-function(t){(N[NM,1]*N[NM,3]*t^(N[NM,3]-
1)+(N[NM,2]*(N[NM,4]+N[NM,5]*t)*t^(N[NM,4]-1)*exp(N[NM,5]*t)))}

#### New Modified Weibull CDF
NMcdf<-function(x,y){
 a<-x[1]
```

```
 b<-x[2]
 t<-x[3]
 g<-x[4]
 l<-x[5]

 F<-c(rep(NA,1))

 F<-(1-exp((-a*y^t)-(b*y^g)*exp(l*y)))

 return(F)

}

NMtest<-NMcdf(x=c(N[NM,1],N[NM,2],N[NM,3], N[NM,4],N[NM,5]),y=Meeker)


###  Bi-Weibull Model ##########################
p0<-matrix(NA, nrow=100, ncol=4, byrow=FALSE,); colnames(p0)<-
c("beta1","beta2","theta1","theta2")
p0[,1]<-c(runif(100, 0.25,5 )) #beta1
p0[,2]<-c(runif(100, 0.25,5 )) #beta2
p0[,3]<-c(runif(100, min(Meeker),max(Meeker))) #theta1
p0[,4]<-c(runif(100, min(Meeker),max(Meeker))) #theta2

GG<-function(x,y,d){
 b1<-x[1]
 b2<-x[2]
 t1<-x[3]
 t2<-x[4]

 F<-c(rep(NA,1))

 F<-sum(log(((b1*(y^(b1-1))*(t1)^(-b1))+(b2*y^(b2-1))*((t2)^(-b2))))*d-
(((y)/t1)^(b1)+(y/t2)^(b2)))

 return(-F)
}

P<-matrix(NA, nrow=100, ncol=5, byrow=FALSE,); colnames(P)<-
c("beta1","beta2","theta1","theta2", "Value")

for (i in 1:100){
 P[i,]=0
```

```
  P[i,1:4]<-optim(par=p0[i,], fn=GG, y=c(Meeker),d=c(Censor), method="L-BFGS-B",
lower=c(0.25, 0.25, 10, 10), upper=c(90, 90, 400, 400)) $par
  P[i,5]<-GG(P[i,1:4],y=c(Meeker),d=c(Censor))
}

PM<-which.min(P[,5])
P[PM,]

#SEB<-optim(par=P[PM,1:4], fn=GG, y=c(Meeker),d=1, method="L-BFGS-B",
lower=c(0.25, 0.25, 10, 10), upper=c(90, 90, 100, 100),hessian=TRUE) $hessian
#seb<-sqrt(diag(solve(SEB)))

####Bi-Weibull Reliability function
BiWeib<-function(t){exp(-1*(((t/(P[PM,3]))^P[PM,1])+((t/(P[PM,4]))^P[PM,2])))}
####Bi-Weibull Hazard function
BiHaz<-function(t){(P[PM,1]*t^(P[PM,1]-
1))/(P[PM,3]^P[PM,1])+(P[PM,2]*t^(P[PM,2]-1))/(P[PM,4]^P[PM,2])}
####Bi-Weibull pdf function
Bipdf<-function(t){((P[PM,1]*t^(P[PM,1]-
1))/(P[PM,3]^P[PM,1])+(P[PM,2]*t^(P[PM,2]-1))/(P[PM,4]^P[PM,2]))*(exp(-
1*(((t/(P[PM,3]))^P[PM,1])+((t/(P[PM,4]))^P[PM,2]))))}

####Bi-Weibull CDF
BWcdf<-function(x,y){
  b1<-x[1]
  b2<-x[2]
  t1<-x[3]
  t2<-x[4]

  F<-c(rep(NA,1))

  F<-1-(exp(-1*(((y/t1)^b1)+((y/(t2))^b2))))

  return(F)

}

BWtest<-BWcdf(x=c(P[PM,1],P[PM,2],P[PM,3],P[PM,4]),y=Meeker)


############################ Tri-Weibull Model
#########################
t0<-matrix(NA, nrow=100, ncol=6, byrow=FALSE,); colnames(t0)<-
c("beta1","beta2","beta3","theta1","theta2","theta3")
```

```
t0[,1]<-c(runif(100, 0.25,5 )) #beta1
t0[,2]<-c(runif(100, 0.25,5 )) #beta2
t0[,3]<-c(runif(100, 0.25,5 )) #beta3
t0[,4]<-c(runif(100, min(Meeker),max(Meeker))) #theta1
t0[,5]<-c(runif(100, min(Meeker),max(Meeker))) #theta2
t0[,6]<-c(runif(100, min(Meeker),max(Meeker))) #theta3


TT<-function(x,y,d){
 b1<-x[1]
 b2<-x[2]
 b3<-x[3]
 t1<-x[4]
 t2<-x[5]
 t3<-x[6]

 F<-c(rep(NA,1))

 F<-sum(log((b1*(y^(b1-1))*(t1)^(-b1))+(b2*y^(b2-1))*((t2)^(-b2))+(b3*y^(b3-
1))*((t3)^(-b3)))*d-(((y)/t1)^(b1)+(y/t2)^(b2)+(y/t3)^(b3)))

 return(-F)
}

Tr<-matrix(NA, nrow=100, ncol=7, byrow=FALSE,); colnames(Tr)<-
c("beta1","beta2","beta3","theta1","theta2","theta3","Value")

for (i in 1:100){
 Tr[i,]=0
 Tr[i,1:6]<-optim(par=t0[i,], fn=TT, y=c(Meeker),d=c(Censor), method="L-BFGS-B",
lower=c(25, 0.25, 0.25, 10, 10, 10), upper=c(124, 90, 90, 1200, 400, 400)) $par
 Tr[i,7]<-TT(Tr[i,1:6],y=c(Meeker),d=c(Censor))
}

TM<-which.min(Tr[,7])
Tr[TM,]

#SET<-optim(par=Tr[TM,1:6], fn=TT, y=c(Meeker),d=1, method="L-BFGS-B",
lower=c(0.25, 0.25, 0.25, 10, 10, 10), upper=c(110, 90, 90, 100, 125,
100),hessian=TRUE) $hessian
#set<-sqrt(diag(solve(SET)))

####Tri-Weibull Reliability function
```

```
TriWeib<-function(t){exp(-
1*(((t/(Tr[TM,4]))^Tr[TM,1])+((t/(Tr[TM,5]))^Tr[TM,2])+((t/(Tr[TM,6]))^Tr[TM,3
]))))}
####Tri-Weibull Hazard function
TriHaz<-function(t){(Tr[TM,1]*t^(Tr[TM,1]-
1))/(Tr[TM,4]^Tr[TM,1])+(Tr[TM,2]*t^(Tr[TM,2]-
1))/(Tr[TM,5]^Tr[TM,2])+(Tr[TM,3]*t^(Tr[TM,3]-1))/(Tr[TM,6]^Tr[TM,3])}
####Bi-Weibull pdf function
Tripdf<-function(t){((Tr[TM,1]*t^(Tr[TM,1]-
1))/(Tr[TM,4]^Tr[TM,1])+(Tr[TM,2]*t^(Tr[TM,2]-
1))/(Tr[TM,5]^Tr[TM,2])+(Tr[TM,3]*t^(Tr[TM,3]-1))/(Tr[TM,6]^Tr[TM,3]))*(exp(-
1*(((t/(Tr[TM,4]))^Tr[TM,1])+((t/(Tr[TM,5]))^Tr[TM,2])+((t/(Tr[TM,6]))^Tr[TM,3
])))))}

####Tri-Weibull CDF
TWcdf<-function(x,y){
 b1<-x[1]
 b2<-x[2]
 b3<-x[3]
 t1<-x[4]
 t2<-x[5]
 t3<-x[6]

 F<-c(rep(NA,1))

 F<-1-(exp(-1*(((y/t1)^b1)+((y/(t2))^b2)+((y/(t3))^b3))))

 return(F)
}

TWtest<-
TWcdf(x=c(Tr[TM,1],Tr[TM,2],Tr[TM,3],Tr[TM,4],Tr[TM,5],Tr[TM,6]),y=Meeker)


############## Quad-Weibull Model ##########################
# q0<-matrix(NA, nrow=100, ncol=8, byrow=FALSE,); colnames(q0)<-
c("beta1","beta2","beta3","beta4","theta1","theta2","theta3","theta4")
# q0[,1]<-c(runif(100, 0.25,5 )) #beta1
# q0[,2]<-c(runif(100, 0.25,5 )) #beta2
# q0[,3]<-c(runif(100, 0.25,5 )) #beta3
# q0[,4]<-c(runif(100, 0.25,5 )) #beta4
# q0[,5]<-c(runif(100, min(Meeker),max(Meeker))) #theta1
# q0[,6]<-c(runif(100, min(Meeker),max(Meeker))) #theta2
# q0[,7]<-c(runif(100, min(Meeker),max(Meeker))) #theta3
```

```
# q0[,8]<-c(runif(100, min(Meeker),max(Meeker))) #theta4
#
# QQ<-function(x,y,d){
#   b1<-x[1]
#   b2<-x[2]
#   b3<-x[3]
#   b4<-x[4]
#   t1<-x[5]
#   t2<-x[6]
#   t3<-x[7]
#   t4<-x[8]
#
#   F<-c(rep(NA,1))
#
#   F<-sum(log((b1*(y^(b1-1))*(t1)^(-b1))+(b2*y^(b2-1))*((t2)^(-b2))+(b3*y^(b3-
1))*((t3)^(-b3))+(b4*y^(b4-1))*((t4)^(-b4)))*d-
(((y)/t1)^(b1)+(y/t2)^(b2)+(y/t3)^(b3)+(y/t4)^(b4)))
#
#   return(-F)
# }
#
# Q<-matrix(NA, nrow=100, ncol=9, byrow=FALSE,); colnames(Q)<-
c("beta1","beta2","beta3","beta4","theta1","theta2","theta3","theta4","Value")
#
# for (i in 1:100){
#   Q[i,]=0
#   Q[i,1:8]<-optim(par=q0[i,], fn=QQ, y=c(Meeker),d=c(Censor), method="L-BFGS-
B", lower=c(0.25, 0.25, 0.25,0.25, 10, 10, 10, 10), upper=c(115, 115, 90, 90, 500, 500,
400, 400)) $par
#   Q[i,9]  <-optim(par=q0[i,], fn=QQ, y=c(Meeker),d=c(Censor), method="L-BFGS-
B", lower=c(0.25, 0.25, 0.25,0.25, 10, 10, 10, 10), upper=c(115, 115, 90, 90, 500, 500,
400, 400)) $value
# }
#
# QM<-which.min(Q[,9])
# Q[QM,]
#
# #SEQ<-optim(par=Q[QM,1:8], fn=QQ, y=c(Meeker),d=1, method="L-BFGS-B",
lower=c(0.25, 0.25, 0.25,0.25, 10, 10, 10, 10), upper=c(110, 100, 90, 90, 100, 100,
125, 100),hessian=TRUE) $hessian
# #seq<-sqrt(diag(SEQ))
#
# ####Quad-Weibull Reliability function
```

```
# QuadWeib<-function(t){exp(-
1*(((t/(Q[QM,5]))^Q[QM,1])+((t/(Q[QM,6]))^Q[QM,2])+((t/(Q[QM,7]))^Q[QM,3])+(
(t/(Q[QM,8]))^Q[QM,4]))))}
# ####Quad-Weibull Hazard function
# QuadHaz<-function(t){(Q[QM,1]*t^(Q[QM,1]-
1))/(Q[QM,5]^Q[QM,1])+(Q[QM,2]*t^(Q[QM,2]-
1))/(Q[QM,6]^Q[QM,2])+(Q[QM,3]*t^(Q[QM,3]-
1))/(Q[QM,7]^Q[QM,3])+(Q[QM,4]*t^(Q[QM,4]-1))/(Q[QM,8]^Q[QM,4])}
# ####Quad-Weibull pdf function
# Quadpdf<-function(t){((Q[QM,1]*t^(Q[QM,1]-
1))/(Q[QM,5]^Q[QM,1])+(Q[QM,2]*t^(Q[QM,2]-
1))/(Q[QM,6]^Q[QM,2])+(Q[QM,3]*t^(Q[QM,3]-
1))/(Q[QM,7]^Q[QM,3])+(Q[QM,4]*t^(Q[QM,4]-1))/(Q[QM,8]^Q[QM,4]))*(exp(-
1*(((t/(Q[QM,5]))^Q[QM,1])+((t/(Q[QM,6]))^Q[QM,2])+((t/(Q[QM,7]))^Q[QM,3])+(
(t/(Q[QM,8]))^Q[QM,4])))))}
#
# ####Quad-Weibull CDF
# QWcdf<-function(x,y){
#   b1<-x[1]
#   b2<-x[2]
#   b3<-x[3]
#   b4<-x[4]
#   t1<-x[5]
#   t2<-x[6]
#   t3<-x[7]
#   t4<-x[8]
#
#   F<-c(rep(NA,1))
#
#   F<-1-(exp(-1*(((y/t1)^b1)+((y/(t2))^b2)+((y/(t3))^b3)+((y/(t4))^b4))))
#
#   return(F)
# }
#
# QWtest<-
QWcdf(x=c(Q[QM,1],Q[QM,2],Q[QM,3],Q[QM,4],Q[QM,5],Q[QM,6],Q[QM,7],Q[QM,8]),
y=Meeker)

############ Single Weibull Model ##########################
Weib<-function(x,y,d){
 b1<-x[1]
 t1<-x[2]
 n<-length(Aarset2)
```

```
 F<-c(rep(NA,1))

 F<-(d*n*(log(b1)-b1*log(t1))+(d*(b1-1))*sum(log(y))-sum(((y)/t1)^b1))

 return(-F)
}

SW<-matrix(NA, nrow=100, ncol=3, byrow=FALSE,); colnames(SW)<-
c("beta","theta","Value")

for (i in 1:100){
 SW[i,]=0
 SW[i,1:2]<-optim(par=c(p0[i,1], p0[i,3]), fn=Weib, y=Meeker, d=Censor,
method="L-BFGS-B", lower=c(0.25, 0.25, 10, 10), upper=c(5, 5, 1000, 1000)) $par
 SW[i,3]<-Weib(SW[i,1:2],y=Meeker,d=Censor)
}
SM<-which.min(SW[,3])

####Single-Weibull Reliability function
UniWeib<-function(t){exp(-1*(((t/(SW[SM,2]))^SW[SM,1])))}
####Single-Weibull Hazard function
UniHaz<-function(t){(SW[SM,1]*t^(SW[SM,1]-1))/(SW[SM,2]^SW[SM,1])}
####Single-Weibull pdf function
UniPdf<-function(t){(SW[SM,1]*t^(SW[SM,1]-1))/(SW[SM,2]^SW[SM,1])*exp(-
1*(((t/(SW[SM,2]))^SW[SM,1])))}

####Single-Weibull CDF
SWcdf<-function(x,y){
 b1<-x[1]
 t1<-x[2]

 F<-c(rep(NA,1))

 F<-1-exp(-1*(((t/(SW[SM,2]))^SW[SM,1])))

 return(F)
}

SWtest<-SWcdf(x=c(SW[SM,1],SW[SM,2]),y=Meeker)


######### Exponentiated modified Weibull extension Model
e0<-matrix(NA, nrow=100, ncol=4, byrow=FALSE,); colnames(p0)<-
c("alpha","lambda","beta","gamma")
```

```
e0[,1]<-c(runif(100, 40,100 )) #alpha
e0[,2]<-c(runif(100, 10^-5,10^-2 )) #lambda
e0[,3]<-c(runif(100, .01,.8)) #beta
e0[,4]<-c(runif(100, 0.25,.625)) #gamma

EE<-function(x,y,z,d){
 alp<-x[1]
 lam<-x[2]
 bet<-x[3]
 gam<-x[4]

 n<-length(Meeker1)

 F<-c(rep(NA,1))

 F<-n*(alp*lam+(1-bet)*log(alp)+log(bet)+log(lam)+log(gam))-
alp*lam*(sum(exp((y/alp)^bet)))+1/(alp^bet)*(sum(y^bet))+(bet-
1)*(sum(log(y)))+(gam-1)*sum(1-exp(lam*alp*(1-exp((y/alp)^bet))))+log(prod(1-
(1-exp(lam*alp*(1-exp((z/alp)^bet)))^gam)))

 return(-F)
}

E<-matrix(NA, nrow=100, ncol=5, byrow=FALSE,); colnames(P)<-
c("alpha","lamba","beta","gamma", "Value")

for (i in 1:100){
 E[i,]=0
 E[i,1:4]<-optim(par=e0[i,], fn=EE, y=c(Meeker1), z=c(Meeker2),d=1,method="L-
BFGS-B",lower=c(40, 10^-4, .01, 0.25), upper=c(100, 10^-2, .8, .625)) $par
 E[i,5]<-EE(E[i,1:4],c(Meeker1),c(Meeker2),1)
}

EM<-which.min(E[,5])
E[EM,]

####EMWE Reliability function
EWeib<-function(t){1-(1-exp(E[EM,2]*E[EM,1]*(1-
exp((t/E[EM,1])^E[EM,3])))))^(E[EM,4])}
####EMWE Hazard function
EHaz<-function(t){(E[EM,2]*E[EM,3]*E[EM,4]*(t/E[EM,1])^(E[EM,3]-
1)*exp((t/E[EM,1])^E[EM,3]+E[EM,2]*E[EM,1]*(1-
exp((t/E[EM,1])^E[EM,3])))))/(((1-exp(E[EM,2]*E[EM,1]*(1-
```

```
exp((t/E[EM,1])^E[EM,3])))^(1-E[EM,4])+exp(E[EM,2]*E[EM,1]*(1-
exp((t/E[EM,1])^E[EM,3]))-1))}
####EMWE pdf
EPdf<-function(t){E[EM,2]*E[EM,3]*E[EM,4]*(t/E[EM,1])^(E[EM,3]-
1)*exp((t/E[EM,1])^E[EM,3]+E[EM,2]*E[EM,1]*(1-exp((t/E[EM,1])^E[EM,3])))*(1-
exp(E[EM,2]*E[EM,1]*(1-exp((t/E[EM,1])^E[EM,3])))^(E[EM,4]-1)}
####EMWE CDF
Ecdf<-function(t){(1-exp(E[EM,2]*E[EM,1]*(1-
exp((t/E[EM,1])^E[EM,3])))^(E[EM,4])}

EMtest<-Ecdf(t=eCDF)

############################ Plot Reliability Functions against K-M
Estimate ############################
KMFit<-survfit(Surv(Meeker, Censor)~1)
plot(KMFit, axes = FALSE, xlab = NA, ylab = NA, xaxs="r", yaxs="r",
ylim=range(c(0,1)))
par(new=TRUE)
curve(NMWeib,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="red",lty=3,xaxs="r", yaxs="r", ylim=range(c(0,1)))
par(new=TRUE)
curve(BiWeib,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,xaxs="r", yaxs="r", ylim=range(c(0,1)))
par(new=TRUE)
curve(TriWeib,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="orange",xaxs="r", yaxs="r", ylim=range(c(0,1)))
par(new=TRUE)
curve(EWeib,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="green",lty=2,xaxs="r", yaxs="r", ylim=range(c(0,1)))
box()
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "t", line = 2)
mtext(side = 2, "R(t)", line = 2.5)
legend("bottomleft", c("Bi-Weibull","Tri-Weibull","NMW",
"EMWE"),lty=c(1,1,3,2),lwd=c(1.5,1.5),col=c("black","orange","red","green"),cex=0.7
5,bty="n",xjust=0.5)

############################ Plot Density Functions against Censored
Histogram ############################
barplot(c(0.004667,0.002,0.002,0.001333,0.002,0.002667),space=0,col="white",
axes=FALSE, xpd=FALSE, xaxs="r",yaxs="r", ylim=range(c(0,0.005)))
```

```
par(new=TRUE)
curve(NMWPdf,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="red",lty=3,xaxs="r", yaxs="r", ylim=range(c(0,0.005)))
par(new=TRUE)
curve(BiPdf,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,xaxs="r", yaxs="r", ylim=range(c(0,0.005)))
par(new=TRUE)
curve(TriPdf,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="orange",xaxs="r", yaxs="r", ylim=range(c(0,0.005)))
# par(new=TRUE)
# curve(EPdf,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="blue",xaxs="r", yaxs="r", ylim=range(c(0,0.005)))
par(new=TRUE)
curve(EPdf,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="green",lty=2,xaxs="r", yaxs="r", ylim=range(c(0,0.005)))
# par(new=TRUE)
# curve(UniPdf,from=0, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab =
NA,col="Purple",xaxs="r", yaxs="r", ylim=range(c(0,0.005)))
box()
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "t", line = 2)
mtext(side = 2, "f(t)", line = 3)
legend("top", c("Bi-Weibull","Tri-Weibull","NMW",
"EMWE"),lty=c(1,1,3,2),lwd=c(1.5,1.5),col=c("black","orange","red","green"),cex=0.7
5,bty="n",xjust=0.5)

barplot(c(.005283,.002791,.003243,.0025,.004444,.013333),space=0,col="white",
axes=FALSE, xpd=FALSE, xaxs="i", ylim=range(c(0,0.015)))
par(new=TRUE)
curve(NMWHaz,from=0.00002, to=max(Meeker),n=1000, axes = FALSE, xlab = NA,
ylab = NA,col="red",lty=3,xaxs="i", yaxs="i", ylim=range(c(0,0.015)))
par(new=TRUE)
curve(BiHaz,from=0.00002, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab
= NA,xaxs="i", yaxs="i", ylim=range(c(0,0.015)))
par(new=TRUE)
curve(TriHaz,from=0.00002, to=max(Meeker),n=1000, axes = FALSE, xlab = NA,
ylab = NA,col="orange",xaxs="i", yaxs="i",ylim=range(c(0,0.015)))
# par(new=TRUE)
# curve(QuadHaz,from=0.00002, to=max(Meeker),n=1000, axes = FALSE, xlab = NA,
ylab = NA,col="blue",xaxs="i", yaxs="i", ylim=range(c(0,0.015)))
par(new=TRUE)
```

```
curve(EHaz,from=0.00002, to=max(Meeker),n=1000, axes = FALSE, xlab = NA, ylab
= NA,col="green",lty=2,xaxs="i", yaxs="i", ylim=range(c(0,0.015)))
# par(new=TRUE)
# curve(UniHaz,from=0.00002, to=max(Meeker),n=1000, axes = FALSE, xlab = NA,
ylab = NA,col="Purple",xaxs="i", yaxs="i", ylim=range(c(0,0.015)))
box()
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "t", line = 2)
mtext(side = 2, "h(t)", line = 3)
legend("top", c("Bi-Weibull","Tri-Weibull","NMW",
"EMWE"),lty=c(1,1,3,2),lwd=c(1.5,1.5),col=c("black","orange","red","green"),cex=0.7
5,bty="n",xjust=0.5)

N[NM,]
P[PM,]

######### Perform K-S Test for both models and compare Against Meeker ECDF

ks.test(RelFreq, NMtest)
ks.test(RelFreq, BWtest)
ks.test(RelFreq, TWtest)
ks.test(RelFreq, QWtest)
ks.test(RelFreq, SWtest)
ks.test(RelFreq, EMtest)

######## Find standard errors for bi-Weibull parameter estimates

h1<-((P[PM,1]*data^(P[PM,1]-1))/(P[PM,3]^(P[PM,1])))
h2<-((P[PM,2]*data^(P[PM,2]-1))/(P[PM,4]^(P[PM,2])))
r1<-(data/(P[PM,3]))^(P[PM,1])
r2<-(data/(P[PM,4]))^(P[PM,2])
hh<-((P[PM,1]*data^(P[PM,1]-1))/(P[PM,3]^(P[PM,1])))+((P[PM,2]*data^(P[PM,2]-
1))/(P[PM,4]^(P[PM,2])))

likeb1ex<-sum(d*(h1/hh*(log(data/P[PM,3])+P[PM,1]^-1)))-
sum(r1*log(data/P[PM,3]))
likeb2ex<-sum(d*(h2/hh*(log(data/P[PM,4])+P[PM,2]^-1)))-
sum(r2*log(data/P[PM,4]))
liket1ex<-sum(-(P[PM,1]/P[PM,3]*d*h1/hh))+sum((P[PM,1]/P[PM,3])*r1)
liket2ex<-sum(-(P[PM,2]/P[PM,4]*d*h2/hh))+sum((P[PM,2]/P[PM,4])*r2)
```

```
likeb1ex
likeb2ex
liket1ex
liket2ex

Hess<-matrix(NA,nrow=4, ncol=4,byrow=FALSE)

Hess[1,1]<-sum((h1*(log(P[PM,3])-log(data))*(log(P[PM,3])-log(data)-
2/P[PM,1])*hh-(h1*(log(data)-log(P[PM,3]+P[PM,1]^-1)))^2)/(hh^2))-
sum(r1*(log(data/P[PM,3]))^2)
Hess[1,2]<-sum((-(h1*(log(data)-log(P[PM,3])+P[PM,1]^-1)*h2*(log(data)-
log(P[PM,4])+P[PM,2]^-1)))/(hh^2))
Hess[1,3]<-sum(((P[PM,1]/P[PM,3])*(-hh*h1*(log(data)-log(P[PM,3])+2*P[PM,1]^-
1)+(P[PM,1]/P[PM,3])*h1^2*(log(data)-log(P[PM,3])+P[PM,1]^-
1)))/(hh^2))+sum((P[PM,1]/P[PM,3])*r1*(log(data/P[PM,3])+P[PM,1]^-1))
Hess[1,4]<-sum((h1*(log(data)-log(P[PM,3])+P[PM,1]^-1)*P[PM,2]*h2*P[PM,4]^-
1)/(hh^2))
Hess[2,1]<-sum((-(h2*(log(data)-log(P[PM,4])+P[PM,2]^-1)*h1*(log(data)-
log(P[PM,3])+P[PM,1]^-1)))/(hh^2))
Hess[2,2]<-sum((h2*(log(P[PM,4])-log(data))*(log(P[PM,4])-log(data)-
2/P[PM,2])*hh-(h2*(log(data)-log(P[PM,4]+P[PM,2]^-1)))^2)/(hh^2))-
sum(r2*(log(data/P[PM,4]))^2)
Hess[2,3]<-sum((h2*(log(data)-log(P[PM,4])+P[PM,2]^-1)*P[PM,1]*h1*P[PM,3]^-
1)/(hh^2))
Hess[2,4]<-sum(((P[PM,2]/P[PM,4])*(-hh*h2*(log(data)-log(P[PM,4])+2*P[PM,2]^-
1)+(P[PM,2]/P[PM,4])*h2^2*(log(data)-log(P[PM,4])+P[PM,2]^-
1)))/(hh^2))+sum((P[PM,2]/P[PM,4])*r2*(log(data/P[PM,4])+P[PM,2]^-1))
Hess[3,1]<-sum(((P[PM,1]/P[PM,3])*(-hh*h1*(log(data)-log(P[PM,3])+2*P[PM,1]^-
1)+(P[PM,1]/P[PM,3])*h1^2*(log(data)-log(P[PM,3])+P[PM,1]^-
1)))/(hh^2))+sum((P[PM,1]/P[PM,3])*r1*(log(data/P[PM,3])+P[PM,1]^-1))
Hess[3,2]<-sum(((P[PM,1]/P[PM,3])*h1*h2*(log(data)-log(P[PM,4])+P[PM,2]^-
1))/(hh^2))
Hess[3,3]<-sum(((P[PM,1]/P[PM,3])*(P[PM,1]+1)*(P[PM,3]^-1)*h1*hh-(-
h1*(P[PM,1]/P[PM,3]))^2)/(hh^2))-sum(((P[PM,1]^2+P[PM,1])*r1)/(P[PM,3]^2))
Hess[3,4]<-sum(((-P[PM,1]/P[PM,3])*h1*(P[PM,2]/P[PM,4])*h2)/(hh^2))
Hess[4,1]<-sum(((P[PM,2]/P[PM,4])*h2*h1*(log(data)-log(P[PM,3])+P[PM,1]^-
1))/(hh^2))
Hess[4,2]<-sum(((P[PM,2]/P[PM,4])*(-hh*h2*(log(data)-log(P[PM,4])+2*P[PM,2]^-
1)+(P[PM,2]/P[PM,4])*h2^2*(log(data)-log(P[PM,4])+P[PM,2]^-
1)))/(hh^2))+sum((P[PM,2]/P[PM,4])*r2*(log(data/P[PM,4])+P[PM,2]^-1))
Hess[4,3]<-sum(((-P[PM,2]/P[PM,4])*h2*(P[PM,1]/P[PM,3])*h1)/(hh^2))
Hess[4,4]<-sum(((P[PM,2]/P[PM,4])*(P[PM,2]+1)*(P[PM,4]^-1)*h2*hh-(-
h2*(P[PM,2]/P[PM,4]))^2)/(hh^2))-sum(((P[PM,2]^2+P[PM,2])*r2)/(P[PM,4]^2))
```

```
SE<-solve(-Hess)
SE

######## Find standard errors for tri-Weibull parameter estimates
data<-c(PMin[S[z]:N[z],u])

h1<-((Tr[TM,1]*data^(Tr[TM,1]-1))/(Tr[TM,4]^(Tr[TM,1])))
h2<-((Tr[TM,2]*data^(Tr[TM,2]-1))/(Tr[TM,5]^(Tr[TM,2])))
h3<-((Tr[TM,3]*data^(Tr[TM,3]-1))/(Tr[TM,6]^(Tr[TM,3])))
r1<-(data/(Tr[TM,4]))^(Tr[TM,1])
r2<-(data/(Tr[TM,5]))^(Tr[TM,2])
r3<-(data/(Tr[TM,6]))^(Tr[TM,3])
hh<-((Tr[TM,1]*data^(Tr[TM,1]-
1))/(Tr[TM,4]^(Tr[TM,1])))+((Tr[TM,2]*data^(Tr[TM,2]-
1))/(Tr[TM,5]^(Tr[TM,2])))+((Tr[TM,3]*data^(Tr[TM,3]-
1))/(Tr[TM,6]^(Tr[TM,3])))

likeb1ex<-sum((h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-1))/hh)-
sum(r1*log(data/Tr[TM,4]))
likeb2ex<-sum((h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-1))/hh)-
sum(r2*log(data/Tr[TM,5]))
likeb3ex<-sum((h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-1))/hh)-
sum(r3*log(data/Tr[TM,6]))
liket1ex<-sum(-(Tr[TM,1]/Tr[TM,4]*h1/hh))+sum((Tr[TM,1]/Tr[TM,4])*r1)
liket2ex<-sum(-(Tr[TM,2]/Tr[TM,5]*h2/hh))+sum((Tr[TM,2]/Tr[TM,5])*r2)
liket3ex<-sum(-(Tr[TM,3]/Tr[TM,6]*h3/hh))+sum((Tr[TM,3]/Tr[TM,6])*r3)

likeb1ex
likeb2ex
likeb3ex
liket1ex
liket2ex
liket3ex

Hess<-matrix(NA,nrow=6, ncol=6,byrow=FALSE)

Hess[1,1]<-sum((h1*(log(Tr[TM,4])-log(data))*(log(Tr[TM,4])-log(data)-
2/Tr[TM,1])*hh-(h1*(log(data)-log(Tr[TM,4]+Tr[TM,1]^-1)))^2)/(hh^2))-
sum(r1*(log(data/Tr[TM,4]))^2)
Hess[1,2]<-sum((-(h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-1)*h2*(log(data)-
log(Tr[TM,5])+Tr[TM,2]^-1)))/(hh^2))
Hess[1,3]<-sum((-(h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-1)*h3*(log(data)-
log(Tr[TM,6])+Tr[TM,3]^-1)))/(hh^2))
```

```
Hess[1,4]<-sum(((Tr[TM,1]/Tr[TM,4])*(-hh*h1*(log(data)-
log(Tr[TM,4])+2*Tr[TM,1]^-1)+(Tr[TM,1]/Tr[TM,4])*h1^2*(log(data)-
log(Tr[TM,4])+Tr[TM,1]^-
1)))/(hh^2))+sum((Tr[TM,1]/Tr[TM,4])*r1*(log(data/Tr[TM,4])+Tr[TM,1]^-1))
Hess[1,5]<-sum((h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-
1)*Tr[TM,2]*h2*Tr[TM,5]^-1)/(hh^2))
Hess[1,6]<-sum((h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-
1)*Tr[TM,3]*h3*Tr[TM,6]^-1)/(hh^2))
Hess[2,1]<-sum((-(h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-1)*h1*(log(data)-
log(Tr[TM,4])+Tr[TM,1]^-1)))/(hh^2))
Hess[2,2]<-sum((h2*(log(Tr[TM,5])-log(data))*(log(Tr[TM,5])-log(data)-
2/Tr[TM,2])*hh-(h2*(log(data)-log(Tr[TM,5]+Tr[TM,2]^-1)))^2)/(hh^2))-
sum(r2*(log(data/Tr[TM,5]))^2)
Hess[2,3]<-sum((-(h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-1)*h3*(log(data)-
log(Tr[TM,6])+Tr[TM,3]^-1)))/(hh^2))
Hess[2,4]<-sum((h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-
1)*Tr[TM,1]*h1*Tr[TM,4]^-1)/(hh^2))
Hess[2,5]<-sum(((Tr[TM,2]/Tr[TM,5])*(-hh*h2*(log(data)-
log(Tr[TM,5])+2*Tr[TM,2]^-1)+(Tr[TM,2]/Tr[TM,5])*h2^2*(log(data)-
log(Tr[TM,5])+Tr[TM,2]^-
1)))/(hh^2))+sum((Tr[TM,2]/Tr[TM,5])*r2*(log(data/Tr[TM,5])+Tr[TM,2]^-1))
Hess[2,6]<-sum((h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-
1)*Tr[TM,3]*h3*Tr[TM,6]^-1)/(hh^2))
Hess[3,1]<-sum((-(h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-1)*h1*(log(data)-
log(Tr[TM,4])+Tr[TM,1]^-1)))/(hh^2))
Hess[3,2]<-sum((-(h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-1)*h2*(log(data)-
log(Tr[TM,5])+Tr[TM,2]^-1)))/(hh^2))
Hess[3,3]<-sum((h3*(log(Tr[TM,6])-log(data))*(log(Tr[TM,6])-log(data)-
2/Tr[TM,3])*hh-(h3*(log(data)-log(Tr[TM,6]+Tr[TM,3]^-1)))^2)/(hh^2))-
sum(r3*(log(data/Tr[TM,6]))^2)
Hess[3,4]<-sum((h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-
1)*Tr[TM,1]*h1*Tr[TM,4]^-1)/(hh^2))
Hess[3,5]<-sum((h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-
1)*Tr[TM,2]*h2*Tr[TM,5]^-1)/(hh^2))
Hess[3,6]<-sum(((Tr[TM,3]/Tr[TM,6])*(-hh*h3*(log(data)-
log(Tr[TM,6])+2*Tr[TM,3]^-1)+(Tr[TM,3]/Tr[TM,6])*h3^2*(log(data)-
log(Tr[TM,6])+Tr[TM,3]^-
1)))/(hh^2))+sum((Tr[TM,3]/Tr[TM,6])*r3*(log(data/Tr[TM,6])+Tr[TM,3]^-1))
Hess[4,1]<-sum(((Tr[TM,1]/Tr[TM,4])*(-hh*h1*(log(data)-
log(Tr[TM,4])+2*Tr[TM,1]^-1)+(Tr[TM,1]/Tr[TM,4])*h1^2*(log(data)-
log(Tr[TM,4])+Tr[TM,1]^-
1)))/(hh^2))+sum((Tr[TM,1]/Tr[TM,4])*r1*(log(data/Tr[TM,4])+Tr[TM,1]^-1))
Hess[4,2]<-sum(((Tr[TM,1]/Tr[TM,4])*h1*h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-
1))/(hh^2))
```

```
Hess[4,3]<-sum(((Tr[TM,1]/Tr[TM,4])*h1*h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-
1))/(hh^2))
Hess[4,4]<-sum(((Tr[TM,1]/Tr[TM,4])*(Tr[TM,1]+1)*(Tr[TM,4]^-1)*h1*hh-(-
h1*(Tr[TM,1]/Tr[TM,4]))^2)/(hh^2))-
sum(((Tr[TM,1]^2+Tr[TM,1])*r1)/(Tr[TM,4]^2))
Hess[4,5]<-sum(((-Tr[TM,1]/Tr[TM,4])*h1*(Tr[TM,2]/Tr[TM,5])*h2)/(hh^2))
Hess[4,6]<-sum(((-Tr[TM,1]/Tr[TM,4])*h1*(Tr[TM,3]/Tr[TM,6])*h3)/(hh^2))
Hess[5,1]<-sum(((Tr[TM,2]/Tr[TM,5])*h2*h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-
1))/(hh^2))
Hess[5,2]<-sum(((Tr[TM,2]/Tr[TM,5])*(-hh*h2*(log(data)-
log(Tr[TM,5])+2*Tr[TM,2]^-1)+(Tr[TM,2]/Tr[TM,5])*h2^2*(log(data)-
log(Tr[TM,5])+Tr[TM,2]^-
1)))/(hh^2))+sum((Tr[TM,2]/Tr[TM,5])*r2*(log(data/Tr[TM,5])+Tr[TM,2]^-1))
Hess[5,3]<-sum(((Tr[TM,2]/Tr[TM,5])*h2*h3*(log(data)-log(Tr[TM,6])+Tr[TM,3]^-
1))/(hh^2))
Hess[5,4]<-sum(((-Tr[TM,2]/Tr[TM,5])*h2*(Tr[TM,1]/Tr[TM,4])*h1)/(hh^2))
Hess[5,5]<-sum(((Tr[TM,2]/Tr[TM,5])*(Tr[TM,2]+1)*(Tr[TM,5]^-1)*h2*hh-(-
h2*(Tr[TM,2]/Tr[TM,5]))^2)/(hh^2))-
sum(((Tr[TM,2]^2+Tr[TM,2])*r2)/(Tr[TM,5]^2))
Hess[5,6]<-sum(((-Tr[TM,2]/Tr[TM,5])*h2*(Tr[TM,3]/Tr[TM,6])*h3)/(hh^2))
Hess[6,1]<-sum(((Tr[TM,3]/Tr[TM,6])*h3*h1*(log(data)-log(Tr[TM,4])+Tr[TM,1]^-
1))/(hh^2))
Hess[6,2]<-sum(((Tr[TM,3]/Tr[TM,6])*h3*h2*(log(data)-log(Tr[TM,5])+Tr[TM,2]^-
1))/(hh^2))
Hess[6,3]<-sum(((Tr[TM,3]/Tr[TM,6])*(-hh*h3*(log(data)-
log(Tr[TM,6])+2*Tr[TM,3]^-1)+(Tr[TM,3]/Tr[TM,6])*h3^2*(log(data)-
log(Tr[TM,6])+Tr[TM,3]^-
1)))/(hh^2))+sum((Tr[TM,3]/Tr[TM,6])*r3*(log(data/Tr[TM,6])+Tr[TM,3]^-1))
Hess[6,4]<-sum(((-Tr[TM,3]/Tr[TM,6])*h3*(Tr[TM,1]/Tr[TM,4])*h1)/(hh^2))
Hess[6,5]<-sum(((-Tr[TM,3]/Tr[TM,6])*h3*(Tr[TM,2]/Tr[TM,5])*h2)/(hh^2))
Hess[6,6]<-sum(((Tr[TM,3]/Tr[TM,6])*(Tr[TM,3]+1)*(Tr[TM,6]^-1)*h3*hh-(-
h3*(Tr[TM,3]/Tr[TM,6]))^2)/(hh^2))-
sum(((Tr[TM,3]^2+Tr[TM,3])*r3)/(Tr[TM,6]^2))

SE<-solve(-Hess)
SE
```

Appendix D: Computer Code Used to Simulate and Analyze Poly-Weibull Failure Observations as Discussed in Chapter V

```
### set.seed(42)
### This code generates observations from n independent, distinct Weibull
### distributions denoting the n independent failure modes in each system.  The
### minimum of these observations is kept as the actual system failure time.  n is
### simulated as either 10,20, 40 or 100.  The failure data thus comes from a Poly-
### Weibull distribution.  The code captures this data into a matrix PMin for
### various combinations of number of systems and number of failure modes

######## CONTROL VARIABLES
a<-5       ## Selects the ath element of vector z for the number of systems
starts<-10  ## The number of starting points in the parameter space
quality<-10 ## Value for the shape parameter theta
f<-10000    ## The number of iterations

z<-as.vector(c(10,20,40,100)) #Observations

######## Initialize matrices of parameters values for each model
SW<-matrix(NA, nrow=starts, ncol=3, byrow=FALSE,)
P<-matrix(NA, nrow=starts, ncol=5, byrow=FALSE,)
Tr<-matrix(NA, nrow=starts, ncol=7, byrow=FALSE,)
PROBs <-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
PROBb <-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
PROBt <-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
OSLM  <-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
OSL5M <-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
OSL10M<-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
OSL20M<-matrix(NA, nrow=f, ncol=5, byrow=TRUE)
MEAN<-matrix(NA, nrow=7, ncol=5, byrow=TRUE)

colnames(P) <-c("beta1","beta2","theta1","theta2", "Value")
colnames(SW)<-c("beta","theta","Value")
colnames(Tr)<-c("beta1","beta2","beta3","theta1","theta2","theta3","Value")
colnames(PROBb) <-c("5-modes","10-modes","20-modes","40-modes","50-modes")
colnames(PROBs) <-c("5-modes","10-modes","20-modes","40-modes","50-modes")
colnames(PROBt) <-c("5-modes","10-modes","20-modes","40-modes","50-modes")
colnames(OSLM)<-c("5-modes","10-modes","20-modes","40-modes","50-modes")
colnames(OSL5M)<-c("5-modes","10-modes","20-modes","40-modes","50-modes")
colnames(OSL10M)<-c("5-modes","10-modes","20-modes","40-modes","50-
modes")
```

```
colnames(OSL20M)<-c("5-modes","10-modes","20-modes","40-modes","50-
modes")
rownames(MEAN)<-c("Weibull", "bi-Weibull", "tri-Weibull","AD(p-
value)","OSL5","OSL10","OSL20")
colnames(MEAN)<-c("5 modes","10 modes","20 modes","40 modes","50 modes")

for (c in 1:f) {

AIC<-matrix(NA, nrow=5, ncol=11, byrow=TRUE); colnames(AIC)<-c("Weib","Bi-
Weib","Tri-Weib","AIC","p-S","p-B","p-T","AD(p-value)", "OSL5","OSL10","OSL20")

############################ Single Weibull Model
#########################
s0<-matrix(NA, nrow=starts, ncol=2, byrow=FALSE,); colnames(s0)<-c("a","b")
s0[,1]<-c(runif(starts, .25,10)) #beta - shape parameter
s0[,2]<-c(runif(starts, 1,1500)) #alpha - scale parameter

############################ Bi-Weibull Model
#########################
p0<-matrix(NA, nrow=starts, ncol=4, byrow=FALSE,); colnames(p0)<-
c("beta1","beta2","theta1","theta2")
p0[,1]<-c(runif(starts, .15,.9 )) #beta1
p0[,2]<-c(runif(starts, 5.0,40 )) #beta2
p0[,3]<-c(runif(starts, 2.5,1000)) #theta1
p0[,4]<-c(runif(starts, 2.5,1000)) #theta2

######## Tri-Weibull Model
t0<-matrix(NA, nrow=starts, ncol=6, byrow=FALSE,)
colnames(t0)<-c("beta1","beta2","beta3","theta1","theta2","theta3")

t0[,1]<-c(runif(starts, .15,.9))  #beta1
t0[,2]<-c(runif(starts, 5,30 )) #beta2
t0[,3]<-c(runif(starts, 1.25,40 )) #beta3
t0[,4]<-c(runif(starts, 2.5,1000)) #theta1
t0[,5]<-c(runif(starts, 2.5,1000)) #theta2
t0[,6]<-c(runif(starts, 2.5,1000)) #theta3

j<-1
i<-1

###### INITIALIZE THE MATRIX THAT THE FAILURE OBSERVATIONS WILL BE
WRITTEN TO
PMin<-matrix(NA, nrow=z[a], ncol=5, byrow=FALSE,)
colnames(PMin)<-c("5-Modes","10-Modes","20-Modes","40-Modes","50-Modes")
```

```
m<-as.vector(c(5, 10, 20, 40, 50))
for (j in 1:5) { # grab an element from the vector m

####### Create Weibull shape and scale parameters for each failure mode
CV<-matrix(NA, nrow=m[j], ncol=2, byrow=FALSE); colnames(CV)<-c("CoV",
"Mean")
NLM<-matrix(NA, nrow=m[j], ncol=2, byrow=FALSE); colnames(NLM)<-c("Betaj",
"Alphaj")

BetaP<-as.matrix(c(1.25,quality))

######### Simulate a coefficient of variation (COV) value
CV[,1]<-c(rbeta(m[j], BetaP[1,], BetaP[2,])*(5.5-.05)+.05)
######### Define Mean
CV[,2]<-c(rbeta(m[j], (1+((5.5-CV[,1])/(5.5-.05)))^2, (1+((CV[,1]-.05)/(5.5-
.05)))^2)*3000)

###############Functions###############
COV<-function(x,y) {betaj<-x[1]

(((((gamma(1+2/betaj))/(gamma(1+1/betaj))^2)-1)^0.5)-y)
 }

######### Generate Solutions #############
for (i in 1:m[j]) {
NLM[i,1]<-uniroot(COV, c(0.1, 65), y=c(CV[i,1])) $root
NLM[i,2]<-CV[i,2]/(gamma(1+1/NLM[i,1]))
 }

b<-c(NLM[,1])
t<-c(NLM[,2])

 for (i in 1:z[a]) {

###Generate possible failure time for each failure mode and select the minimum
value
###as the observed system failure time

   PMin[i,j]<-min(rweibull(m[j], b, t))

 }
 data<-sort(PMin[,j])
```

```
######### Compare Fit of the Weibull, Poly-Weibull and New Modified Weibull
Models

############################ Single Weibull Model
#########################
 Weib<-function(x,y){b1<-x[1]; t1<-x[2]

 F<-c(rep(NA,1))

 F<-(z[a]*(log(b1)-b1*log(t1))+(b1-1)*sum(log(y))-sum(((y)/t1)^b1))

 return(-F)
 }
 SW<-matrix(NA, nrow=starts, ncol=3, byrow=FALSE,); colnames(SW)<-
c("beta","theta","Value")

 for (i in 1:starts){
  SW[i,]=0
  SW[i,1:2]<-nlminb(c(s0[i,]), Weib, y=c(PMin[,j]), lower=c(.25,.5),
upper=c(40,750)) $par
  SW[i,3]<-Weib(SW[i,1:2],c(PMin[,j])) #value of the negative log likelihood
function
 }
 SM<-which.min(SW[,3])### Select the minimum value of the 10 solutions

###Perform the Anderson-Darling Goodness of Fit test for each data set
AD<-matrix(NA, nrow=z[a], ncol=1, byrow=TRUE)
for (i in 1:z[a]) {

### Weibull Modified test statistic
AD[i,1]<-(log(1-exp(-1*(data[i]/SW[SM,2])^(SW[SM,1])))-((data[z[a]+1-
i])/SW[SM,2])^(SW[SM,1]))*((1-2*i)/z[a])
}
 ADstar<-(sum(AD[,1])-z[a])*(1+0.2/sqrt(z[a]))
 OSL<-0
 if(ADstar<.474) OSL<-.25
 if(.474<=ADstar&&ADstar<= .637) OSL<-(.1-.25)/(.637-.474)*(ADstar-.474)+.25
 if(.637<=ADstar&&ADstar<= .757) OSL<-(.05-.1)/(.757-.637)*(ADstar-.637)+.10
 if(.757<=ADstar&&ADstar<= .877) OSL<-(.025-.05)/(.877-.757)*(ADstar-.757)+.05
 if(.877<=ADstar&&ADstar<=1.038) OSL<-(.01-.025)/(1.038-.877)*(ADstar-
.877)+.025
 if(ADstar>1.038) OSL<-.01

######## Perform Bernoulli test for each data set
```

```
 OSL5 <-ifelse(OSL>0.05, 1, 0)
 OSL10<-ifelse(OSL>0.10, 1, 0)
 OSL20<-ifelse(OSL>0.20, 1, 0)

####### Compute the corrected AIC value for the single Weibull
AIC[j,1]<-2*2-2*(-SW[SM,3])+((2*2*(2+1))/(z[a]-2-1))

 ####### Bi-Weibull Model
 GG<-function(x,y,d){
  b1<-x[1]; b2<-x[2]; t1<-x[3]; t2<-x[4]

  F<-c(rep(NA,1))

  F<-sum(log(((b1*(y^(b1-1))*(t1)^(-b1))+(b2*y^(b2-1))*((t2)^(-b2))))*d-
(((y)/t1)^(b1)+(y/t2)^(b2)))
  return(-F)
 }

 P<-matrix(NA, nrow=starts, ncol=5, byrow=FALSE,); colnames(P)<-
c("beta1","beta2","theta1","theta2", "Value")

 for (i in 1:starts){
  P[i,]=0
  P[i,1:4]<-nlminb(c(p0[i,1:4]), GG, y=c(PMin[,j]),d=1, lower=c(.25,.25,.5,.5),
upper=c(70,100,1000,1000)) $par
  P[i,5]<-GG(P[i,1:4],c(PMin[,j]),1)
 }
 PM<-which.min(P[,5]);P[PM,]

 AIC[j,2]<-2*4-2*(-P[PM,5]) +((2*4*(4+1))/(z[a]-4-1))

 ############################ Tri-Weibull Model
#########################
 TT<-function(x,y,d){
  b1<-x[1]; b2<-x[2]; b3<-x[3]; t1<-x[4]; t2<-x[5]; t3<-x[6]

  F<-c(rep(NA,1))

  F<-sum(log((b1*(y^(b1-1))*(t1)^(-b1))+(b2*y^(b2-1))*((t2)^(-b2))+(b3*y^(b3-
1))*((t3)^(-b3)))*d-(((y)/t1)^(b1)+(y/t2)^(b2)+(y/t3)^(b3)))

  return(-F)
}
```

```
 Tr<-matrix(NA, nrow=starts, ncol=7, byrow=FALSE,); colnames(Tr)<-
c("beta1","beta2","beta3","theta1","theta2","theta3","Value")

 for (i in 1:starts){
  Tr[i,]=0
  Tr[i,1:6]<-nlminb(c(t0[i,1:6]), TT,y=c(PMin[,j]),d=1, lower=c(.15,.15,.15,.5,.5,.5),
upper=c(100,100,100,1000,1000,1000)) $par
  Tr[i,7]<-TT(Tr[i,1:6],c(PMin[,j]),1)
 }
 TM<-which.min(Tr[,7]);Tr[TM,]

 AIC[j,3] <-2*6-2*(-Tr[TM,7])+((2*6*(6+1))/(z[a]-6-1))
 AIC[j,4] <-which.min(AIC[j,1:3])
 AIC[j,5] <-exp((AIC[j,AIC[j,4]]-AIC[j,1])/2)
 AIC[j,6] <-exp((AIC[j,AIC[j,4]]-AIC[j,2])/2)
 AIC[j,7] <-exp((AIC[j,AIC[j,4]]-AIC[j,3])/2)
 AIC[j,8] <-OSL
 AIC[j,9] <-OSL5
 AIC[j,10] <-OSL10
 AIC[j,11] <-OSL20
}

 PROBs  [c,] <-AIC[,5 ]
 PROBb  [c,] <-AIC[,6 ]
 PROBt  [c,] <-AIC[,7 ]
 OSLM   [c,] <-AIC[,8 ]
 OSL5M  [c,] <-AIC[,9 ]
 OSL10M [c,] <-AIC[,10]
 OSL20M [c,] <-AIC[,11]
}

MEAN[1,]<-c(mean(PROBs [,1]),mean(PROBs  [,2]),mean(PROBs  [,3]),mean(PROBs
[,4]),mean(PROBs  [,5]))
MEAN[2,]<-c(mean(PROBb [,1]),mean(PROBb  [,2]),mean(PROBb
[,3]),mean(PROBb  [,4]),mean(PROBb  [,5]))
MEAN[3,]<-c(mean(PROBt [,1]),mean(PROBt  [,2]),mean(PROBt  [,3]),mean(PROBt
[,4]),mean(PROBt  [,5]))
MEAN[4,]<-c(mean(OSLM  [,1]),mean(OSLM   [,2]),mean(OSLM   [,3]),mean(OSLM
[,4]),mean(OSLM   [,5]))
MEAN[5,]<-c(sum(OSL5M  [,1])/f,sum(OSL5M [,2])/f,sum(OSL5M
[,3])/f,sum(OSL5M [,4])/f,sum(OSL5M [,5])/f)
MEAN[6,]<-c(sum(OSL10M
[,1])/f,sum(OSL10M[,2])/f,sum(OSL10M[,3])/f,sum(OSL10M[,4])/f,sum(OSL10M[,5
])/f)
```

```
MEAN[7,]<-c(sum(OSL20M
[,1])/f,sum(OSL20M[,2])/f,sum(OSL20M[,3])/f,sum(OSL20M[,4])/f,sum(OSL20M[,5
])/f)

###Diagnostic plots to verify assess the code

qqPlot(c(PMin[,j]), dist="weibull", shape=SW[SM,1], scale=SW[SM,2])
CCCC<-sort(c(PMin[,j]))
Ppos<-pp(CCCC, a=0)
quant<-(log(1/(1-Ppos)))
plot(log10(c(CCCC))~log10(quant))

UniPdf<-function(t){(SW[SM,1]*t^(SW[SM,1]-1))/(SW[SM,2]^SW[SM,1])*exp(-
1*(((t/(SW[SM,2]))^SW[SM,1])))}

BiPdf<-function(t){((P[PM,1]*t^(P[PM,1]-
1))/(P[PM,3]^P[PM,1])+(P[PM,2]*t^(P[PM,2]-1))/(P[PM,4]^P[PM,2]))*(exp(-
1*(((t/(P[PM,3]))^P[PM,1])+((t/(P[PM,4]))^P[PM,2])))))}

TriPdf<-function(t){((Tr[TM,1]*t^(Tr[TM,1]-
1))/(Tr[TM,4]^Tr[TM,1])+(Tr[TM,2]*t^(Tr[TM,2]-
1))/(Tr[TM,5]^Tr[TM,2])+(Tr[TM,3]*t^(Tr[TM,3]-1))/(Tr[TM,6]^Tr[TM,3]))*(exp(-
1*(((t/(Tr[TM,4]))^Tr[TM,1])+((t/(Tr[TM,5]))^Tr[TM,2])+((t/(Tr[TM,6]))^Tr[TM,3
])))))}

hist(c(PMin[,j]))#, breaks=z[a]/4)
par(new=TRUE)
curve(UniPdf, from=2.5-min(PMin[,j]),to=max(PMin[,j]), col="red",n=2000, xlab="",
ylab="", axes=FALSE, xaxs="r", yaxs="r")
par(new=TRUE)
curve(BiPdf, from=2.5-min(PMin[,j]),to=max(PMin[,j]), col="blue",n=2000, xlab="",
ylab="", axes=FALSE, xaxs="r", yaxs="r")
par(new=TRUE)
curve(TriPdf, from=2.5-min(PMin[,j]),to=max(PMin[,j]), col="green",n=2000,
xlab="", ylab="", axes=FALSE, xaxs="r", yaxs="r")
KMFit<-survfit(Surv(c(PMin[,j]), c(rep(1,z[a])))~1)

#####Single-Weibull Reliability function
UniWeib<-function(time){exp(-1*(((((time))/(SW[SM,2]))^SW[SM,1])))}

#####Bi-Weibull Reliability function
BiWeib<-function(time)
{exp(-1*((((time)/(P[PM,3]))^P[PM,1])+(((time)/(P[PM,4]))^P[PM,2])))}
```

```
#####Tri-Weibull Reliability function
TriWeib<-function(time)
{exp(-1*((((time)/(Tr[TM,4]))^Tr[TM,1])+(((time)/(Tr[TM,5]))^Tr[TM,2])+
(((time)/(Tr[TM,6]))^Tr[TM,3])))}

######## Plot Reliability Functions against K-M Estimate
plot(KMFit, axes = FALSE, xlab = NA, ylab = NA, xaxs="r", yaxs="r",
ylim=range(c(0,1)))
par(new=TRUE)
curve(UniWeib,from=0,to=max(PMin[,j]),n=100,axes=FALSE,xlab=NA,ylab=NA,col="
red",lty=3,xaxs="r",yaxs="r",ylim=range(c(0,1)))
par(new=TRUE)
curve(BiWeib,from=0,to=max(PMin[,j]),n=100,axes=FALSE,xlab=NA,ylab=NA,lty=1,
xaxs="r",yaxs="r", ylim=range(c(0,1)))
par(new=TRUE)
curve(TriWeib,from=0, to=max(PMin[,j]),n=100, axes = FALSE, xlab = NA,
ylab=NA,col="orange",xaxs="r", yaxs="r", ylim=range(c(0,1)))

axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1,"t", line = 2.0)
mtext(side = 2, "R(t)", line = 2.5)


## Generate Histograms
g<-c(0,.05,.1,.15,.2,.25,.3,.35,.4,.45,.5,.55,.6,.65,.7,.75,.8,.85,.9,.95,1)
setwd("C:/Users/Jason/Desktop")
#
#
tiff(filename = "5by100.tiff",
    width = 1800, height = 2700, units = "px", pointsize = 12,
    compression = c("none"),bg = "white", res = 300)

par(mfrow = c(3,1),oma=rep(4, 4), mar=rep(1, 4) )

hist(PROBs[,1], col=rgb(1,0,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), ylim=c(0,f), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
```

```
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Weibull", line = .5)

hist(PROBb[,1], col=rgb(0,1,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), ylim=c(0,f), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Bi-Weibull", line = .5)

hist(PROBt[,1],col=rgb(0,0,1,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1),ylim=c(0,f), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Tri-Weibull", line = .5)

mtext( expression("AIC"[corrected]*" (5,100,Low)" ), outer = TRUE)
dev.off(2)
#
#
tiff(filename = "10by100.tiff",
   width = 1800, height = 2700, units = "px", pointsize = 12,
   compression = c("none"),bg = "white", res = 300)

par(mfrow = c(3,1),oma=rep(4, 4), mar=rep(1, 4) )

hist(PROBs[,2], col=rgb(1,0,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Weibull", line = .5)
```

```
hist(PROBb[,2], col=rgb(0,1,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Bi-Weibull", line = .5)

hist(PROBt[,2], col=rgb(0,0,1,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Tri-Weibull", line = .5)

mtext( expression("AIC"[corrected]*" (10,100,Low)" ), outer = TRUE)
dev.off(2)
#
#
tiff(filename = "20by100.tiff",
   width = 1800, height = 2700, units = "px", pointsize = 12,
   compression = c("none"),bg = "white", res = 300)

par(mfrow = c(3,1),oma=rep(4, 4), mar=rep(1, 4) )

hist(PROBs[,3],col=rgb(1,0,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Weibull", line = .5)

hist(PROBb[,3], col=rgb(0,1,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
```

```
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Bi-Weibull", line = .5)

hist(PROBt[,3],col=rgb(0,0,1,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Tri-Weibull", line = .5)

mtext( expression("AIC"[corrected]*" (20,100,Low)" ), outer = TRUE)
dev.off(2)
#
#
tiff(filename = "40by100.tiff",
    width = 1800, height = 2700, units = "px", pointsize = 12,
    compression = c("none"),bg = "white", res = 300)

par(mfrow = c(3,1),oma=rep(4, 4), mar=rep(1, 4) )

hist(PROBs[,4],col=rgb(1,0,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Weibull", line = .5)

hist(PROBb[,4], col=rgb(0,1,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
```

```
mtext(side = 4, "Bi-Weibull", line = .5)

hist(PROBt[,4], col=rgb(0,0,1,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Tri-Weibull", line = .5)

mtext( expression("AIC"[corrected]*" (40,100,Low)" ), outer = TRUE)
dev.off(2)
#
#
tiff(filename = "50by100.tiff",
    width = 1800, height = 2700, units = "px", pointsize = 12,
    compression = c("none"),bg = "white", res = 300)

par(mfrow = c(3,1),oma=rep(4, 4), mar=rep(1, 4) )

hist(PROBs[,5], col=rgb(1,0,0,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1),breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Weibull", line = .5)

hist(PROBb[,5], col=rgb(0,1,0,.25),axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Bi-Weibull", line = .5)

hist(PROBt[,5], col=rgb(0,0,1,.25), axes = FALSE,main="", xlab = NA, ylab = NA,
xaxs="r", yaxs="r", xlim=c(0,1), breaks=g)
```

```
axis(side = 1, tck = -.015, labels = NA)
axis(side = 2, tck = -.015, labels = NA)
axis(side = 1, lwd = 0, line = -.6)
axis(side = 2, lwd = 0, line = -.6, las = 1)
mtext(side = 1, "", line = 2.0)
mtext(side = 2, "frequency", line = 2.5)
mtext(side = 4, "Tri-Weibull", line = .5)

mtext( expression("AIC"[corrected]*" (50,100,Low)" ), outer = TRUE)
dev.off(2)
#
#
```

Vita

Major Jason K. Freels graduated from Clinton High School in Clinton, Tennessee. He entered undergraduate studies at Auburn University in Auburn, Alabama where he graduated with a Bachelor of Science in Aerospace Engineering in December of 2000. He was commisioned through AFROTC Detachment 005 at Auburn University.

His first assignment was to Hill AFB, Utah in February 2001 where he served as a depot engineer. In August 2004, he was assigned to the Air Force Institute of Technology to pursue a graduate degree. Upon earning a Master of Science in Materials Science and Engineering in August of 2006, he was assigned to the Air Force Research Laboratory in the Materials and Manufacturing Directorate to conduct research in non-destructing evaluation and inspection. In June 2008 he was assigned as an Air Force Research Laboratory liason to the Aeronautical Systems Center. In August of 2010 he entered the Graduate School of Engineering and Management, Air Force Institute of Technology. Upon graduation he will be assigned to the Air Force Institue of Technology as an assistant professor of Systems Engineering.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
20-12-2013	PhD Dissertation	Sep 2010 - Dec 2013

4. TITLE AND SUBTITLE

Modeling Reliability Growth in Accelerated Stress Testing

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5. AUTHOR(S)

Freels, Jason K, Major USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENS-DS-13-D-02

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Intentionally left blank

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Qualitative accelerated test methods improve system reliability by identifying and removing initial design flaws. However, schedule and cost constraints often preclude sufficient testing to generate a meaningful reliability estimate from the data obtained in these tests. In this dissertation a modified accelerated life test is proposed to assess the likelihood of attaining a reliability requirement based on tests of early system prototypes. Assuming each prototype contains an unknown number of independent competing failure modes whose respective times to occurrence are governed by a distinct Weibull law, the observed failure data from this qualitative test are shown to follow a poly-Weibull distribution. However, using an agent-based Monte Carlo simulation, it is shown that for typical products subjected to qualitative testing, the failure observations result from a homogenous subset of the total number of latent failure modes and the failure data can be adequately modeled with a Weibull distribution. Thus, the projected system reliability after implementing corrective action to remove one or more failure modes can be estimated using established quantitative accelerated test data analysis methods. Our results suggest that a significant cost and time savings may be realized using the proposed method to signal the need to reassess a product's design or reallocate test resources to avoid unnecessary maintenance or redesigns. Further, the proposed approach allows a significant reduction in the test time and sample size required to estimate the risk of meeting a reliability requirement over current quantitative accelerated life test techniques. Additional contributions include a numerical and analytical procedure for obtaining the maximum likelihood parameter estimates and observed Fisher information matrix components for the generalized poly-Weibull distribution. Using this procedure, we show that the poly-Weibull distribution outperforms the best-fit modified Weibull alternatives in the literature with respect to their fit of reference data sets for which the hazard rate functions are non-monotone.

15. SUBJECT TERMS

Reliability Growth Planning Models, Quantitative Accelerated Life Testing, Qualitative Accelerated Life Testing, Competing Risks Analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Joseph J. Pignatiello, AFIT/ENS
U	U	U	UU	240	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636 ext 4311; joseph.pigntiello@afit.edu